JOURNAL *of* EDUCATIONAL RESEARCH

Volume XXXIII NOVEMBER, 1939 Number 3

A MINIMUM READING-SPELLING VOCABULARY FOR REMEDIAL WORK

GEORGE SPACHE
Friends Seminary, New York City

Editor's note: With the development of more technical information relative to elementary education and the professionalization of teaching, there has come a demand for better instructional materials and aids to teaching. The author attempts to offer a tentative average grade placement for words in the primary grades.

THE number of basic reading and spelling vocabularies offered the remedial teacher has multiplied rapidly in recent years. A few of the more frequently used reading vocabularies include a list of "the 1811 words highly suitable for use in all forms of reading material in grades 1, 2 and 3" (3), a graded vocabulary of 2000 words for primary grades, arranged in eight levels (8), a list of the "most important 150 words for earliest book reading" (9), the most important 453 words in the first-grade vocabulary (11), a "basic sight vocabulary of 220 words" (1) as well as a compilation of "the 20,000 words found most frequently and widely in general reading of children and young people" (10). Constant additions to this array are being made by extensive counts of current readers (5).

Core vocabularies in spelling include those drawn from the vocabulary of children's letters, adult writing vocabularies, words frequently misspelled by children, association studies and compilations drawn from several of these sources. In all, Foran (2) notes 46 spelling vocabularies of these various types. Compilations of spoken vocabularies include those of the vocabulary of children before entering the first grade (6) and the commonest 1000 words in the spoken vocabulary of children (7).

Differences in the criteria of selection, in the sources and even in the methods of tabulating roots and derived word forms make it extremely difficult for the remedial teacher to choose a basic or minimum vocabulary upon which to base her work. A further complication arises from the fact that children, recommended for remedial teaching in reading or spelling,

are frequently found to be disabled in both fields. Since reading vocabularies for any particular level, however, are distinctly larger than spelling vocabularies for the same grade, it becomes necessary to employ two different basic lists. Moreover, the spelling list may include words not present in the reading vocabulary, and hence, not basic to immediate reading needs. Thus, when attempting remedial teaching in reading and spelling simultaneously, the teacher is faced with the necessity of checking every "basic" reading word against the "basic spelling vocabulary" and every "basic" spelling word against the "basic reading vocabulary."

In an attempt to combat these difficulties, the writer has compared the three most recent comprehensive reading and spelling lists. The spelling vocabulary was that published by A. I. Gates as "A List of Spelling Difficulties in 3876 Words" (4). The list includes the most frequently recurring words in twenty-five different widely used spelling textbooks, and state and large-city spelling lists. It undoubtedly includes the 3876 words most frequently taught in American classrooms. Since many of the words were placed at different grade levels in the various sources, Gates computed an average grade placement by determining the mean grade placement in the texts and state or city lists.

The reading vocabularies employed were those of Stone, "A Graded Vocabulary for Primary Reading" (8) and the word-count of the vocabularies of recently published preprimers of Hayward and Ordway (5). The Stone list is composed of eight reading levels termed respectively, the pre-book stage, easy preprimers and first half of easy primers, difficult preprimers, second half of easy primers and first half of more difficult primers, latter half of difficult primers and easy first readers, more difficult first-readers, easier second grade readers, more difficult second-grade readers and third-reader material. These levels are enumerated since the writer's list follows the same divisions.

Stone's placement of the 2000 words in reading levels was determined by (1) the range and order of appearance of words as new words in primers (and preprimers of the same series) and in first, second and third readers, a total of 99 books; (2) the rank in the Gates 1926 and 1935 lists (3); (3) the rank in the Wheeler-Howell list of the 453 most important words for first-grade reading (11); and (4) appearance in the spoken vocabulary of children up to and including six years of age (6, 7).

The Hayward and Ordway compilation (5) was used since it used seven pre-primers of later publication date as well as eight of the twelve

employed by Stone. Forty-three words were found in the Hayward-Ordway list not present in Stone, 36 of these being proper names found in one or two preprimers. The remaining seven words, baggage, cheep, dining, engineer, mew-mew, pullman and shoo, were also peculiar to one of the 15 preprimers. Thus, their list offers no additions to the Stone vocabulary.

One possible correction of the Stone list that may be derived from the Hayward-Ordway study is in the placement of words in the various reading levels. Twelve words appear in two or more of the preprimers studied by Hayward-Ordway that Stone would place in level IV or above, i.e., words suitable for difficult primers or easy first readers or above. The appearance of these words as new words in primers and first readers, as noted by Stone, and their appearance in preprimers by Hayward-Ordway are tabulated in Table I.

TABLE I

READING-LEVEL PLACEMENT OF 12 WORDS IN STONE AND HAYWARD-ORDWAY

	Stone			Hayward-Ordway
	Level	No. of Primers	No. of 1st Readers	No. of Preprimers
airplane	IV	5	6	2
airplanes	..			2
boat	V	4	8	2
boats	..			1
cap	V	5	9	2
caps	..			1
circus	V	2	9	2
dear	IV	9	13	3
hello	VI	5	4	2
noise	V	4	9	9
puppy	V	5	6	3
ready	IV	9	10	2
something	IV	11	11	4
store	IV	12	8	5

The writer has computed from Stone's list that the words in levels I-II occurred in at least 14 of the 29 primers as new words and in less than 20 of the 27 first readers as new words. The median number of books in which a word in levels I-II appeared as a new word is 3.7 preprimers, 24.6 primers and 4.1 first readers. Words in level III appeared as new words in a median number of 11.6 primers and 9.8 first readers. Words in level IV appeared as new words in a median number of 6.4 primers and 10.6 first readers in Stone's list. Words in level V appeared as new

words in a median number of 2.9 primers and 8.3 first readers. It appears that the words noise, something, and store might be placed at a lower level than IV or V. Because of their frequent appearance in preprimers, the words may be placed at level III, i.e., words suitable for difficult preprimers, second half of easy primers and first half of more difficult primers.

This tentative re-grading of three words merely demonstrates that, despite the inclusion of seven later preprimers than those employed by Stone, the reading level placement of only three words is questioned when the Hayward-Ordway list is compared with Stone's.

The following minimum reading-spelling vocabulary was determined by checking Stone's entire list against the Gates average spelling grade placement. Words not present in the Gates list were thus eliminated from the Stone list. Therefore, the minimum reading-spelling vocabulary is not serviceable as a list against which words may be checked to determine whether they are basic to reading needs alone. It is not serviceable as a basic spelling list alone. For these two needs we recommend the Stone and Gates vocabularies, respectively. The minimum reading-spelling vocabulary is a basic list of words suitable for simultaneous remedial teaching in both spelling and reading. It represents a list of words basic to the reading-spelling vocabulary rather than a list basic to either reading or spelling alone.

Its values to the remedial teacher are obvious. In those instances in which a reading disability is accompanied by one in spelling or vice versa, and they are many, the list represents a convenient, accurate, and economical point of departure.

The writer has attempted to retain the distinctive feature of the Stone vocabulary, namely, the division of words into reading levels. Therefore, in his list the words are retained at the reading level for which they were judged most suitable by Stone, with the exception of the three words re-graded by the use of the Hayward-Ordway list.

Next to each word is noted the mean spelling grade placement as given by Gates. Since the minimum vocabulary is largely intended for the primary level of teaching, words in which the spelling grade placement is 5.0 or higher are not included. The general agreement that spelling is to be delayed until the second grade is evidenced by the fact that no words are placed in the first grade in the Gates list.

The Stone list would count some roots and their derived forms as single words, e.g., work, worked, is counted as a single word. The mini-

mum reading-spelling vocabulary counts such words separately. This results in the minimum vocabulary being greater in number of words in some levels than the Stone list from which it was derived. Thus, in levels I-II our list includes 148 of the 150 words placed at that level by Stone. Separate counting of roots and derived forms results in the grand total of 161 words at levels I-II. Similarly, the minimum vocabulary includes 93 of the 100 words in level III of Stone's list, or a grand total of 102 words. At level IV it includes 136 of the 150 possible, and a total of 148. At level V it includes 162 of the 210 possible, and a total of 178. At level VI it includes 251 of the 350 possible, and a total of 259. At level VII, the minimum reading-spelling vocabulary includes 205 of the 458 of Stone's list. Counting roots and derived forms separately brings the total to 210. At level VIII, the minimum vocabulary includes 162 of the 555 words placed there by Stone. The total at this level is 167. Of the entire 2000 words in the Stone vocabulary, the minimum vocabulary includes 1157. Counting roots and derived forms separately, the total number of words in the minimum vocabulary is 1225.

The minimum reading-spelling vocabulary includes 216 of the 220 words comprising the Dolch basic sight vocabulary, which is composed of the words (exclusive of nouns) common to the Kindergarten-Union (6) and Wheeler-Howell (11) lists and the first 500 words of the Gates vocabulary (3). Our minimum list includes 469 of the first 500 words of the Gates vocabulary (3) and 315 of the 453 words judged most important for first grade by Wheeler and Howell (11).

To facilitate the use and understanding of the following lists, the details of their derivation may be repeated:

Levels I-II—Most Important Words for Beginning Reading

These words were placed at the lowest levels of the reading-spelling vocabulary because of their frequent occurrence as new words in the 12 preprimers and 29 primers studied by Stone (9). On the average, they occur as new words in approximately one-third of the preprimers and 84 percent* of the primers. They occur as new words in only 15 percent of the 27 first readers.

Level III—Easy Primer Material

On the average, words of this level occurred as new words in 40 percent of the primers and 36 percent of the first readers. Stone (9) de-

*These and similar figures derived by the writer by computing the median number of books in which each word appeared.

scribes them as words characteristic of the vocabulary of difficult pre-primers, the second half of easy primers and the first half of the more difficult primers.

Level IV—Difficult Primer and Easy First-Reader Material

These words were presented, on the average, as new words in 22 percent of the primers and 39 percent of the first readers. They are characteristic of the vocabulary of difficult primers and easy first readers.

Level V—Difficult First-Reader Material

The average words of this level occurred as new words in 10 percent of the primers, 30 percent of the first readers, 28 percent of the second readers and 15 percent of the third readers. In other words, they may be described as the vocabulary commonly found in the more difficult first reader.

Level VI—Easy Second-Reader Material

Words of this level are characteristic of easier second grade reading materials. The average word appeared as a new word in 8 percent of the first readers, 43 percent of the second readers and 26 percent of the third readers reviewed by Stone (9).

Level VII—Difficult Second-Reader Material

Words of this level are commonly found in the more difficult second grade reader. They appeared as new words, on the average, in 5 percent of the first readers, 28 percent of the second, and 29 percent of the third readers.

Level VIII—Easy Third-Reader Material

In Stone's group of readers, the average word of this level appeared as a new word in 4 percent of the first readers, 15 percent of the second and 39 percent of the third readers. They are the vocabulary commonly found in the easy third grade readers.

The Minimum Reading-Spelling Vocabulary

LEVELS I-II—MOST IMPORTANT WORDS FOR BEGINNING READING

after	2.7	baby	2.3	board	3.6	can	2.3
all	2.2	back	2.4	boy	2.4	cat	2.2
am	2.3	ball	2.2	brown	4.9	chair	2.9
and	2.3	bed	2.4	but	2.4	children	4.1
apple	2.4	big	2.2	by	2.4	come	2.4
are	2.3	bird	2.7	call	2.4	could	3.0
at	2.3	black	2.9	called	3.1	cow	2.8
away	2.3	blue	2.6	came	2.4	day	2.3

did	2.3	him	2.2	one	2.6	the	2.3
do	2.2	his	2.2	out	2.3	them	2.3
doing	2.4	home	2.3	pig	2.3	then	2.6
dog	2.3	house	2.8	play	3.1	there	2.4
doll	2.4	houses	3.6	pretty	3.8	they	2.3
door	2.6	I	2.5	put	2.6	this	3.0
down	2.3	in	2.2	rabbit	3.0	three	2.3
eat	2.3	into	3.4	ran	2.7	to	2.4
eating	4.0	is	2.3	read	3.0	too	3.3
egg	2.9	it	2.3	reading	2.8	tree	2.3
fast	2.6	jump	3.6	red	3.0	two	2.7
father	2.7	laugh	3.9	ride	3.6	under	2.4
find	2.4	like	2.7	run	2.4	up	2.9
for	2.4	little	2.3	said	2.6	us	2.4
from	2.3	look	2.5	saw	3.2	very	2.9
fun	2.9	looked	3.1	say	2.3	want	2.9
gave	2.3	made	3.0	says	3.3	wanted	3.8
get	2.3	make	2.7	school	2.4	was	2.5
girl	2.5	making	3.0	see	2.3	water	2.9
give	2.1	man	2.4	she	2.3	way	2.9
go	2.4	may	2.6	sing	3.1	we	2.5
good	2.4	me	2.5	sleep	2.6	went	2.6
green	3.7	milk	2.9	sleeping	4.8	were	2.4
had	2.3	morning	3.0	so	3.1	what	2.3
has	2.2	mother	2.6	some	2.8	where	2.8
have	2.2	my	2.4	soon	2.8	white	2.8
he	2.2	no	3.6	stop	2.8	who	2.8
help	2.9	not	2.6	stopped	4.3	will	2.4
helped	4.0	now	2.7	table	2.6	with	2.5
hen	3.1	of	2.3	take	2.6	yes	2.8
her	2.3	oh	4.4	thank	4.1	you	2.7
here	2.4	on	2.6	that	2.3	your	2.7

LEVEL III—EASY PRIMER MATERIAL

again	3.2	book	2.3	Christmas	3.1	fall	2.4
an	2.1	box	2.6	coming	2.7	falling	3.8
as	2.2	boxes	3.0	cry	3.2	farm	4.1
asleep	4.8	bread	3.1	draw	3.7	farmer	4.3
ate	3.2	bring	2.5	drink	3.3	feet	2.8
barn	3.8	butter	3.4	duck	3.0	first	2.6
be	2.4	buy	2.7	ear	2.4	five	2.6
being	4.8	cake	2.7	every	2.8	flew	3.8
bear	2.9	car	2.7	eye	2.7	fly	2.6
bell	2.4	catch	3.5	eyes	3.0	found	3.3

four	3.1	live	2.8	pet	3.8	story	2.7
funny	3.9	lives	3.3	picture	4.5	sun	2.4
good-by	4.9	lived	4.1	please	3.8	tail	3.7
got	2.6	long	3.1	pull	3.0	thing	3.0
guess	3.9	Mr.	4.1	pulled	3.5	time	2.8
hand	2.3	must	3.1	right	2.9	took	3.2
happy	4.1	name	2.6	roll	2.9	train	3.3
hat	2.2	nest	2.8	rolled	2.6	wake	2.9
head	2.7	new	2.6	round	2.8	walk	3.1
hear	2.3	night	2.8	sat	2.8	walked	4.0
hill	3.1	noise	3.7	shall	3.2	when	2.4
hop	2.0	old	2.3	sheep	3.3	wind	3.7
horse	3.4	once	3.2	sit	2.6	window	4.1
how	2.6	open	3.4	sitting	4.5	yellow	3.9
kitten	3.1	our	2.5	something	3.7		
let	2.3	party	3.9	store	2.7		

LEVEL IV—DIFFICULT PRIMER AND EASY FIRST-READER MATERIAL

about	3.3	dear	2.8	goes	3.0	light	3.6
airplane	3.5	dinner	3.0	grass	2.9	lost	2.7
always	3.6	does	2.9	gray	4.8	many	2.7
arm	2.9	dress	2.5	great	4.4	met	3.0
ask	2.7	dressed	4.5	ground	3.5	mice	4.1
asked	3.1	drum	3.5	grow	2.8	Miss	3.9
bag	3.8	each	3.6	hair	2.9	mouse	3.1
balloon	3.5	early	3.7	hard	2.6	month	3.4
bee	2.9	engine	4.2	hay	2.8	Mrs.	4.3
been	2.4	face	2.8	heard	3.2	much	2.8
before	3.4	far	2.7	hide	3.1	near	3.8
began	3.8	fat	2.3	high	3.4	never	2.7
birthday	3.7	feed	2.4	hit	2.9	next	3.0
breakfast	4.4	fell	2.8	hole	3.4	nuts	3.1
bump	3.5	fine	3.4	hour	4.0	off	4.1
candy	2.8	fire	3.2	hungry	4.3	only	2.7
carry	3.2	fish	2.6	if	2.2	or	3.1
chicken	4.5	flower	4.0	its	2.9	other	3.1
chickens	3.8	flowers	3.3	just	3.1	over	3.1
coat	3.1	foot	2.7	keep	2.7	paint	3.5
cold	2.3	fox	3.8	kite	3.5	painted	3.5
color	4.1	friend	4.5	know	3.1	pick	3.3
corn	3.2	game	2.8	last	2.3	picked	3.3
cried	3.1	garden	3.4	lay	2.7	place	3.4
cup	2.9	glad	2.6	leg	2.7	plant	3.8
cut	2.5	goat	3.5	letter	2.8	rain	3.0

ready	3.1	sleepy	3.8	ten	3.1	why	2.8
ring	3.1	snow	2.8	their	2.8	wish	2.7
road	2.7	song	3.4	think	2.9	wished	3.3
robin	3.5	spot	3.3	today	2.6	wishing	4.8
sang	3.4	squirrel	4.8	top	2.8	wood	3.9
Santa Claus	2.9	stand	3.1	wagon	3.4	woods	3.8
seed	3.1	stay	2.8	warm	2.7	work	2.8
show	2.8	street	3.1	wash	3.1	worked	3.6
showed	4.0	talk	3.2	washed	3.3	would	3.6
sister	2.8	teacher	3.2	which	3.2	yard	3.1
sled	2.6	tell	2.5	while	3.9		

LEVEL V—DIFFICULT FIRST-READER MATERIAL

across	4.2	climb	4.3	hid	2.5	own	3.8
afraid	3.4	clock	3.6	himself	4.4	owned	4.6
along	3.7	colt	4.0	hold	2.6	pail	3.6
another	3.8	cook	2.9	hot	2.6	pan	2.8
answer	4.3	count	3.2	ice	2.4	paper	2.7
any	2.9	country	4.4	I'll	4.3	park	3.9
bake	4.3	cross	4.1	kick	4.1	paw	3.8
band	3.7	dance	4.5	kind	3.0	peas	4.0
bark	2.6	dark	2.7	kindly	4.8	pen	3.1
basket	3.7	dish	4.0	king	3.8	penny	4.8
because	3.3	dishes	4.6	leaf	3.6	picnic	4.8
behind	3.9	don't	3.6	learn	4.0	pie	3.8
best	2.6	drank	3.5	leaves	3.1	pocket	3.8
better	3.1	drop	2.7	left	3.0	pony	3.6
bigger	3.9	dropped	3.6	limb	4.5	poor	3.0
blow	3.1	eight	4.2	lion	3.1	puppy	4.0
boat	2.7	end	3.0	longer	3.5	race	4.2
bright	4.1	everything	4.1	love	3.6	room	2.9
broke	3.1	field	4.7	lunch	4.1	scratch	4.3
brother	2.7	fill	3.3	march	4.5	seen	3.3
brush	4.6	filled	3.8	men	2.8	seven	3.4
bug	2.5	flies	4.8	monkey	4.3	shake	4.1
cage	4.0	floor	3.7	more	2.8	shine	4.6
calf	3.7	food	3.8	nail	4.3	shining	4.5
cap	2.4	full	3.4	nine	3.8	shoe	2.9
care	3.6	gate	2.4	nose	3.1	shoes	2.8
caught	3.7	gone	2.6	nothing	3.9	should	3.9
circus	4.9	goose	4.6	number	4.4	side	3.1
clap	4.0	grew	3.8	orange	4.1	six	3.1
clean	3.2	happen	4.4	owe	4.6	sky	3.8

slide	3.8	string	3.4	town	2.6	wall	3.5
slow	3.6	summer	3.1	trick	4.3	watch	4.1
small	3.4	supper	3.6	tried	3.7	watched	4.1
smell	4.1	swap	3.6	trunk	4.8	wear	3.4
soft	3.4	swing	3.9	try	2.8	weed	4.8
sometime	4.0	swinging	3.6	trying	4.0	well	2.4
spring	3.2	teeth	3.8	turkey	4.1	wet	3.1
start	3.5	than	2.9	turn	3.6	wing	3.5
started	4.5	these	3.7	turned	3.8	winter	2.5
step	3.4	thought	3.4	until	4.3	without	4.7
steps	4.1	threw	2.8	upon	3.4	woke	4.2
stick	2.8	till	2.9	visit	4.0	woman	4.0
still	3.1	tire	3.6	wait	3.6	worm	3.8
stood	3.3	tired	3.9	waited	3.8		
stories	3.3	told	2.7	waiting	4.0		

LEVEL VI—EASY SECOND-READER MATERIAL

afternoon	4.2	building	4.3	drive	3.2	frog	4.0
ago	2.9	bush	4.5	drove	3.4	front	4.6
air	3.0	busy	4.3	dry	2.8	frost	3.7
almost	4.6	buzz	4.8	dug	2.5	fruit	4.8
alone	3.7	cabbage	4.6	dust	2.9	fur	3.1
anything	3.6	cannot	4.3	easy	3.7	gather	4.8
aunt	2.8	can't	2.9	enough	4.6	gold	3.1
bad	2.3	carried	4.8	even	3.4	grade	4.3
bank	4.2	cent	2.9	ever	2.8	gun	3.6
bath	4.1	cheek	3.5	fairy	3.8	half	3.3
bean	4.1	child	4.2	family	4.6	hammer	4.1
beat	3.7	cloth	4.3	feather	4.5	hang	3.1
begin	3.7	clothes	4.8	fed	2.8	hate	3.6
beside	4.5	cooky	2.8	feel	3.6	hated	4.3
bill	3.5	cool	3.8	felt	3.3	heel	3.6
bite	2.8	corner	4.1	fence	3.3	held	3.3
block	3.8	cover	3.2	few	4.1	hello	4.8
both	3.1	covered	3.8	fight	3.7	hope	2.0
bottom	4.1	crack	4.1	finger	3.6	hundred	3.3
bought	3.9	cracker	4.5	finish	4.2	hung	3.6
bow	4.1	crawl	4.9	finished	4.0	hunt	3.2
branches	4.5	cream	3.4	flag	2.7	hurried	4.1
brave	4.6	crow	3.6	flour	4.3	hurry	4.2
breast	4.8	deep	3.1	follow	4.1	hurt	3.6
broken	4.9	dig	2.8	forget	3.1	inside	3.6
brought	3.3	dime	4.3	forgot	3.4	invite	4.6
build	3.8	done	2.6	fresh	4.7	it's	4.3

jar	3.9	older	2.6	sea	4.4	sure	4.1
joy	4.4	outside	4.8	seat	3.9	sweet	3.1
kept	3.5	page	3.8	seem	4.6	tall	2.6
kitchen	4.7	part	3.3	sell	3.7	taste	4.9
knew	3.2	path	4.1	send	2.6	teach	3.9
lady	4.2	people	4.6	set	3.6	tent	4.1
laid	4.1	piece	4.3	shell	4.2	those	3.9
lamp	3.9	pin	2.6	shook	3.3	thread	4.2
land	3.3	pink	4.0	short	4.0	throw	3.6
large	3.9	plate	4.2	shut	3.2	tie	3.3
late	2.6	plum	3.6	sick	2.8	tied	3.3
lazy	3.8	porch	4.8	sight	4.4	tiny	4.1
lie	3.2	post	3.9	silver	4.2	toe	3.3
lift	3.8	pot	2.8	skip	4.3	tomorrow	4.4
line	3.9	push	3.3	smaller	3.6	tonight	3.4
listen	4.2	question	4.9	smoke	3.3	track	4.2
log	3.6	quick	4.0	smoking	4.2	true	3.9
matter	4.1	rake	4.8	snap	4.5	tub	3.6
maybe	4.6	rang	3.3	sold	3.6	uncle	3.8
mean	3.3	rat	2.3	somebody	3.7	use	3.6
meat	4.1	reach	4.0	sorry	3.7	used	3.6
meet	3.3	reached	4.3	sound	3.2	vine	4.1
might	4.1	real	4.6	soup	4.5	wave	4.7
mind	3.6	rest	3.1	south	3.8	week	2.7
minute	4.1	rich	3.7	spade	3.1	wheat	4.8
money	4.6	river	3.1	spin	4.5	wheel	3.8
moon	3.4	rock	3.6	spread	4.7	whistle	4.4
most	2.7	rode	2.8	star	2.9	whole	3.9
mountain	4.1	roof	4.1	sting	3.8	wide	3.3
move	3.2	root	3.5	stockings	3.8	wild	3.4
mud	2.6	rope	2.8	stone	4.1	word	3.1
music	4.2	rose	2.8	straw	3.9	wore	4.8
myself	4.4	rub	2.8	strong	4.3	world	4.4
neck	3.1	sad	4.2	stuck	4.2	write	3.3
need	4.5	safe	4.8	such	3.8	wrote	3.4
north	3.6	same	2.6	sugar	3.7	year	2.5
o'clock	4.2	sand	3.1	suit	3.8	yesterday	3.6
often	4.0	Saturday	3.9	Sunday	3.2		

LEVEL VII—DIFFICULT SECOND-READER MATERIAL

above	3.6	arrow	4.6	became	4.3	bone	4.4
add	3.5	awake	3.7	beet	4.8	bonnet	4.5
ahead	4.3	ax	3.7	belong	3.7	born	4.6
also	4.4	bare	3.9	below	4.6	bottle	3.9
among	4.8	barrel	4.4	between	4.4	break	3.7
anywhere	3.5	bat	3.8	body	4.0	brick	4.8

bridge	4.8	free	4.7	match	4.6	salt	3.8
broom	4.5	geese	4.8	meal	4.1	saucer	4.8
bucket	4.0	giant	4.6	melt	4.0	save	4.6
built	4.1	gift	4.6	middle	4.2	saving	4.8
burn	3.2	given	4.4	mile	4.1	sent	2.9
burst	4.6	glass	3.3	mill	2.8	sheet	4.3
button	3.9	gloves	4.8	mine	2.8	shelf	4.8
camp	4.9	God	3.1	miss	3.2	ship	3.2
card	3.7	grapes	3.8	Monday	4.5	shop	3.8
careful	4.4	grind	4.7	month	4.6	shot	3.2
cart	4.8	growl	3.6	nearly	3.8	sir	4.1
cattle	4.1	hall	4.0	neat	4.1	skate	3.1
cave	4.8	handle	4.5	net	4.6	skin	3.7
center	4.9	haven't	4.7	nickel	4.7	slept	4.4
change	4.6	heart	4.1	nobody	4.8	snake	4.3
changed	4.6	heavy	4.2	none	4.3	speak	3.7
choose	4.8	hoe	4.3	noon	3.2	spider	4.3
church	4.1	hog	2.8	nurse	4.6	stairs	4.2
class	3.4	hood	4.4	oak	4.1	stamp	3.9
close	3.8	hoop	3.6	oats	4.1	steal	3.3
closed	4.4	ill	4.3	oil	4.2	strike	4.0
coal	4.1	I'm	3.8	oven	4.7	swept	3.8
coffee	4.3	Indian	4.0	paid	4.7	tag	3.8
comb	3.8	instead	4.5	pair	3.9	tea	3.9
cost	3.3	kill	2.9	pass	4.8	tear	3.6
cousin	4.5	killed	4.3	pat	3.0	tease	4.0
crumb	4.6	kiss	3.6	pay	3.1	that's	4.6
daisy	4.5	knee	4.0	pigeon	4.6	thick	4.6
danger	4.1	ladder	4.8	pile	4.0	thin	4.8
deer	4.1	lake	3.9	pillow	4.9	third	4.3
desk	3.7	lamb	4.0	pine	4.5	though	4.9
didn't	3.6	lap	4.2	pipe	4.0	tight	4.4
die	3.3	later	4.8	plow	4.9	tool	3.3
died	3.7	lead	3.8	poem	4.8	touch	4.3
dove	2.3	led	3.4	point	4.6	understand	4.8
drag	3.6	lesson	3.7	pole	4.1	wade	4.5
dream	4.0	life	3.9	pond	3.6	welcome	4.7
drew	4.2	load	3.8	pour	4.6	west	3.6
earth	4.7	loaf	4.8	pray	4.3	whom	4.6
east	3.6	lose	3.7	quiet	4.9	won	4.7
empty	3.9	lots	3.6	quite	4.5	won't	3.6
everybody	4.2	loud	4.6	rag	3.5	wool	4.8
fair	3.6	lovely	4.8	raise	4.3	wrong	4.2
farther	3.8	low	3.4	ribbon	4.2	yet	3.3
flame	4.8	lump	4.4	ripe	3.8	young	4.6
fork	3.6	mail	4.2	row	4.3		
fourth	4.6	mamma	2.5	sack	4.2		

LEVEL VIII—EASY THIRD-READER MATERIAL

able	4.8	edge	4.9	meant	4.8	smooth	4.4
act	4.6	fasten	4.7	mix	3.6	soap	3.8
alive	3.5	February	4.8	muddy	2.8	spend	3.6
America	4.0	fit	4.8	nap	3.6	spill	3.1
anybody	4.8	fix	3.1	needle	4.1	spoil	3.8
ashes	4.1	float	4.1	noisy	3.3	spoiled	3.8
awful	4.9	flock	4.1	nor	4.1	spoke	4.4
backward	4.8	fold	4.1	offer	4.5	stable	4.3
bead	3.8	fond	4.7	ought	4.1	stalk	4.8
become	4.6	fool	3.5	pack	4.1	starve	4.6
bend	3.5	fought	4.9	papa	2.3	steep	4.3
bent	3.3	fountain	4.5	pasture	4.8	stole	3.5
bitter	4.5	Friday	4.3	patch	3.8	stool	4.6
blame	4.2	gas	4.9	peach	4.7	stove	3.2
boil	4.6	grand	4.2	pear	4.3	struck	4.6
brass	4.5	hatch	3.6	pencil	3.6	study	3.6
bunch	4.2	hatched	3.3	pound	4.0	studying	4.2
careless	4.9	holiday	4.8	pump	4.3	tame	3.6
case	4.4	honest	4.6	railroad	4.9	thousand	4.6
chain	4.3	hooks	4.3	rather	4.2	Thursday	4.4
chase	4.2	hotel	4.6	rattle	4.6	ticket	4.7
cherries	4.3	hour	3.8	rice	4.3	tin	3.1
cherry	3.6	inch	3.7	rise	4.7	tip	3.3
chin	3.3	iron	3.8	rob	4.0	tooth	3.8
chop	4.4	jaw	4.8	rubber	3.3	tore	3.6
clear	4.7	jelly	4.9	rug	3.4	torn	4.6
cloak	4.8	juice	4.8	saddle	4.1	trade	4.2
clover	4.3	kettle	4.3	sailor	4.5	tramp	4.3
club	4.3	key	4.4	sew	4.1	trip	4.4
coast	4.3	knife	3.7	shade	3.9	truth	4.5
cotton	4.8	knot	4.5	sharp	3.3	twelve	4.8
cupboard	4.8	lace	4.9	shed	3.8	twice	4.4
dare	4.9	lame	3.6	shirt	4.3	unless	4.6
darling	4.5	lean	4.4	shoot	2.8	vacation	4.2
date	4.7	leather	4.5	sidewalk	4.6	whip	3.1
dead	3.3	less	4.3	silk	4.6	whose	3.8
December	4.3	lick	4.0	since	4.3	wipe	3.6
dirt	3.0	lightning	4.6	sink	3.6	wire	4.1
dirty	3.6	lip	2.8	skirt	4.4	worse	4.3
doctor	4.0	lock	3.6	slap	4.4	worth	4.9
dollar	4.0	mark	3.3	slip	4.0	wouldn't	3.8
dull	3.7	mat	4.6	slipped	3.8		

BIBLIOGRAPHY

1. Dolch, E. W. "A Basic Sight Vocabulary," *Elementary School Journal*, XXXVI (February, 1936), pp. 456-460.
2. Foran, Thomas George. *The Psychology and Teaching of Spelling* (Washington: Catholic Education Press, 1934), 234 pp.
3. Gates, A. I. *A Reading Vocabulary for the Primary Grades* (New York: Teachers College, Columbia University, 1935), 29 pp.
4. Gates, A. I. *A List of Spelling Difficulties in 3876 Words* (New York: Teachers College, Columbia University, 1937), 166 pp.
5. Hayward, W. George, and Ordway, Nancy M. "Vocabularies of Recently Published Preprimers," *Elementary School Journal*, XXXVII (April, 1937), pp. 608-617.
6. Child Study Committee of the International Kindergarten Union. *A Study of the Vocabulary of Children before Entering the First Grade*. Washington: International Kindergarten Union, 1928.
7. National Society for the Study of Education, *Report of the National Committee on Reading, Twenty-fourth Yearbook, Part I*. (Bloomington, Illinois: Public School Publishing Company, 1925), pp. 186-193.
8. Stone, Clarence R. *A Graded Vocabulary for Primary Reading* (St. Louis: Webster Publishing Company, 1936), 61 pp.
9. Stone, Clarence R. *Better Primary Reading* (St. Louis: Webster Publishing Company, 1936), 536 pp.
10. Thorndike, Edward L. *A Teacher's Word Book of the Twenty Thousand Words Found Most Frequently and Widely in General Reading of Children and Young People* (New York: Teachers College, Columbia University, 1932), 182 pp.
11. Wheeler, H. E., and Howell, Emma A. "A First Grade Vocabulary Study," *Elementary School Journal*, XXXI (September, 1930), pp. 52-60.

FUNCTIONAL TEXTBOOKS IN EDUCATIONAL PSYCHOLOGY

JAMES LYNCH
South River High School, South River, N. J.

Editor's note: One of the very important problems of current professional education is the relation of theory and practice. No very satisfactory solution to this problem seems to have arisen. Some persons seem to feel that the solution will be found in more technique courses; others appear to feel that there should be more emphasis upon principles. The author writes upon this very important problem.

"IF you open a book on physics," wrote Woodworth in his *Contemporary Schools of Psychology*, "you do not expect to find a chapter on the wheelbarrow and another on pumps, nor even on the radio, though these useful human objects may be brought in as illustrations of fundamental principles."[1]

Until quite recently the idea behind this statement of Woodworth aptly described the content and structure of textbooks in all fields. The textbook in science, for example, defined its subject in such a way as to clearly distinguish it from all other subjects. Then, the subject was analyzed into its elements, and the elements were arranged according to some logical scheme somewhat as follows: I. Organic Matter, a) Solids, b) Liquids, c) Gases; II. Inorganic Matter, a) Plants, b) Animals, etc. Chapter headings, in the English texts, were named after the parts of speech and arranged in the same order: The Noun, The Pronoun, The Verb, The Adjective, The Adverb, The Conjunction, The Preposition, and The Interjection.

Educators and students of the learning process, however, began to feel that this type of textbook, both in content and in organization, was too preoccupied with definitions, general principles, and systematic classification; too disposed to deal with subject matter without sufficient reference to its role in everyday life. They argued that compartmentalization obscured the practical nature of knowledge, and pointed out further that logical arrangements were the arrangements of the expert—the person who had already learned—and were, therefore, not suitable for the beginner or learner. What the learner needs, they claimed, is material of immediate practicability—material that will function in his own life.

Consequently, there came into prominence the so-called "functional" text, presenting not "Solids," "Liquids," and "Gases," in the field of

[1] Woodworth, R. S. *Contemporary Schools of Psychology* (New York: The Ronald Press, 1931), p. 15.

science, but "The Air We Breathe," "The Water We Drink," "The Food We Eat;" not "Addition," "Equations," "Radicals," etc., in Algebra, but "The Cost of Building a Boat," "The Taxes Your Father Pays," "Excavating Your Cellar," and "Buying Supplies"; not "The Noun," "The Pronoun," "The Adjective," etc., in the English textbook, but "The Library," "Writing a Letter," "Giving a Talk," and "Making Social Calls." Usually, no formal definitions are given; terminology, it is assumed, will come through experience and understanding. Attempts are made, in various ways, to avoid artificial barriers between the different subject-matter areas. Thus, it is common to find the steam-engine considered in a geography text; the thermometer, the automobile, and latitude and longtitude taken up in an algebra book; and equations, formulae, and graphs treated as topics in a civics or in an economics book.

A survey of textbooks in the field of educational psychology reveals a development parallel to this tendency toward functional content and organization.

Educational psychology, as an application of the principles of general psychology to learning situations, used to be patterned on the textbooks of general psychology. There were chapters on: The Nervous System and Learning, Sensation and Learning, Perception and Learning, Attention and Learning, Emotion and Learning, Memory and Learning, Imagination and Learning, and Thinking and Learning. Much of the material was as deeply concerned with learning situations of the general environment as it was with the learning that takes place in school activities.

After the publication of Thorndike's *Educational Psychology*, however, the organization and content of subsequent texts began to change. In the first place, there occurred an almost universal switch from the "nervous system, sensation, perception, memory, imagination, thinking" sequence to the "original nature, laws of learning, individual differences" type of arrangement. Secondly, more material utilizing the results of experiments in learning was introduced: conditioned responses, trial and error learning, learning curves, the distribution of practice, transfer, the effects of fatigue, etc. Some authors added chapters on intelligence testing and on the psychology of the school-subjects, reading, spelling, handwriting, arithmetic, and the social studies.

Even so, there continued a widespread dissatisfaction on the part of teachers and prospective teachers with the texts that came from the press. In spite of the fact that they were based largely upon the work

of an educational psychologist who was recognized as a leader in his field, these texts differed greatly from one another; some to such an extent that a student might conceivably master one of them and not know anything about much of the contents of the others. Remmers and Knight, in 1922, after an attempted analysis of the contents of five texts that deal directly with educational psychology, found "a decided divergence in point of view, organization, and content; so much so, in fact, that a projected comparison of them had to be abandoned."[2] Four years later, Dean A. Worcester, on investigating twelve texts in educational psychology, reported that the topic of instinct was the only item discussed under the same name in all the books analyzed.[3] Criticisms of educational psychology textbooks made in 1927 by one hundred and one teachers of psychology in professional schools for teachers, listed the following weaknesses:

Too theoretical.
Insufficient illustrations.
Ignore classroom situations.
Theories not applicable to actual teaching.

Robinson's study[4] (from which these criticisms were taken), on the basis of conclusions drawn as to the nature of the service which psychology should render to teachers for elementary schools, recommended that educational psychology be presented as it bears upon teaching situations rather than as a distinct branch of academic subject matter.

As if in answer to these comments, there has appeared, within the last three or four years, the so-called functional educational psychology textbook. Here, the subject is approached through the *current* experimental literature on learning, the aim being to treat only those topics which have an up-to-date experimental support and a bearing on modern classroom procedure. While the experimental results of the work carried on by Cattell, Binet, Ebbinghaus, James, Rice, Book, Watson, Woodworth, Thorndike, and others, are not neglected or slighted, the major emphasis is upon accounts of the more recent investigations published

[2] Remmers, H. H., and Knight, F. B. "The Teaching of Educational Psychology in the United States," *Journal of Educational Psychology*, XIII, No. 7, p. 406.

[3] Worcester, D. A. "The Wide Diversities of Practice in First Courses in Educational Psychology," *Journal of Educational Psychology*, XVIII (January, 1927), pp. 11-17.

[4] Robinson, C. L. *Psychology and the Preparation of the Teacher for the Elementary School*. New York: Bureau of Publications, Teachers College, Columbia University, 1930.

in present-day research bulletins and monographs and in the latest issues of such periodicals as the *Journal of Educational Psychology*, the *Journal of Educational Research*, the *Journal of Applied Psychology*, the *Psychological Review*, the *Elementary School Journal*, and the *Review of Educational Research*.

A second objective of the newer texts is to center interest upon the dynamic aspects of the individual learner himself. Hence, they do not lean upon that type of experimental study which has been too much concerned with averages, general laws, and typical tendencies, with its attendant undervaluation of individual differences in needs, capacities and interests. Since they are written from the standpoint that the teacher in the practical work of the classroom does not meet with generalized minds, but with all kinds of concrete differences in reading habits, methods of study, and ways of learning; with children who excel in one subject, are very poor in another, and fair in the rest; with cases of intellectual efficiency coupled with social deficiency in the same individual; and so on; data which are valid only for large groups are not accorded the place they used to hold heretofore. Instead, there is an increased stress upon researches in the fields of guidance techniques, the gifted child, the subnormal child, the maladjusted child, differentiating instruction, study habits, and mental hygiene.

A third objective, really implied in the second, is to deal principally with the facts of learning as they are found under actual classroom conditions. Too many of the findings of experimental psychology which had been published for application to school situations, it is held, were obtained in artificial laboratory set-ups; too many of the experiments emphasized a few years ago, dealt with simple, isolated processes as though the primary concern of psychology were solely with the relationships between stimulus and response elements. What teachers and students of teaching need, on the contrary, is more experimentation with practical learning problems—with how children learn in school. With this in mind, the authors of the various functional texts have extended the scope of educational psychology by including the more complex forms of learning which occur outside the laboratory. In contrast with the practices of their predecessors, therefore, they place a high value on the affective, evaluative, and aesthetic factors in learning, and bring in a wealth of material on such topics as: incentives and interests, the measurement of attitudes, personality traits, and appreciation in literature and in the arts.

Such in briefest outline are some of the innovations that have guided the construction of the functional texts in educational psychology. That they have made up for many of the shortcomings of the older texts cannot be doubted, but, it may be asked, have they not gone too far in certain directions? In the matters of exposition, arrangement, and selection of content, for instance, are not the writers of these newer texts too extreme in their reaction against definition, classification, systematization, and general principles?

At about the same time that educational psychology texts were becoming more practical, there was an increase in the number of experimental investigations devoted to the study of educational terminology—of the terms that teachers and prospective teachers need to know. Rugg, in 1930, became convinced from a study of this problem, that "educational terms or concepts are likely to cause more difficulty for students [of teaching] than any other single factor."[5] Writing in 1937, Carter V. Good placed first in a list of minimum essentials for the development of any scientific field: "a uniform and technical terminology through which to express concepts and to communicate information."[6] Indeed, so important has the Seventeenth National Council of Phi Delta Kappa (1938) deemed the problem of terminology that it has undertaken as a project a dictionary of educational terms. Yet, in the so-called functional educational psychology texts, such words as: educational age, Dalton plan, memory, psychoses, psychoneuroses, neurasthenia, etc., are used without the least attempt at definition.

It seems that the newer authors have misinterpreted the shortcoming of the older type of text in the matter of definition. No doubt a real defect of the older text was its preoccupation with definitions "as such"—definitions without the support of concrete facts. Authors of the newer texts, however, in their desire to supply the student with a foundation of factual knowledge, have rushed directly to the current experimental literature on learning without considering the importance of definitions. In their haste, it seems, they have overlooked the fact that when terms are not defined, the material printed in the most up-to-date periodical is just about as meaningless as the most abstract definition written in the middle ages.

A similar misinterpretation of the limitations of the older texts is

[5] Rugg, E. U. "What Educational Terminology Should a Beginning Teacher Know?" *Educational Administration and Supervision*, XVI (March, 1930), p. 187.
[6] Good, C. V. Quoted in *The Phi Delta Kappan*, XX (November, 1937), p. 82.

revealed in the neglect of system or an orderly arrangement of topics by the newer writers. In spite of the fact that the trouble with the older books was not system per se, but *dead* system, i.e., system isolated from vital experiences, the reaction against the older books seems to have taken the form of a rejection of all system as well as of the type of content systematized. Thus, while the newer texts deal with concrete classroom facts, many technical terms and concepts are used long before they are taken up in detail and explained. Coefficients of correlation and P. E.'s, for example, are brought in time after time in connection with the current experimental literature embodied in the earlier parts of the book while they are not explained until they are dealt with as topics or as subtopics in a later chapter on statistics. The result is a medley of miscellaneous, unrelated facts held together only by the covers of the book and by a title which would be not a little difficult to justify. In fact, some of the newer educational psychologies are no more textbooks of educational psychology than is the *Readers Digest* a textbook on modern history.

The experimental literature on the problems of learning is unquestionably an excellent approach to a functional educational psychology, but, in the case of the books that have appeared thus far, there is much confusion and difference of opinion with respect to the selection of content from this literature. Eurich and Carroll,[7] for example, excuse an inadequate treatment of the problems of emotional adjustment on the ground that not enough is known about the emotions. Similarly, claiming a lack of available knowledge about the physiological and neurological bases of learning, they do not include chapters on these topics. Individual differences, however, they make a major topic of their book, two whole chapters being devoted to it alone. Davis,[8] on the other hand, who also confines his book to a consideration of only those topics for which there is an adequate experimental support, has a chapter on the neurological and physiological bases of learning, but no chapter on individual differences. Stroud[9] starts with the instinct hypothesis, a topic not treated in the Eurich and Carroll book, and proceeds to the discussion of emotion, the physiological basis of mental life, etc. He has no chapter on individual differences.

[7] Eurich, A. C., and Carroll, H. A. *Educational Psychology.* Boston: D. C. Heath & Company, 1935.

[8] Davis, R. A. *Psychology of Learning: A Textbook in Educational Psychology.* New York: McGraw-Hill Book Company, 1935.

[9] Stroud, J. B. *Educational Psychology.* New York: The Macmillan Co., 1935.

Moreover, the experimental work of the Gestalt school of psychology receives little or no attention. This is all the more surprising not only because there is scarcely a current periodical which does not contain some of the results of the Gestalt experiments on learning, but also because these experiments are carried on from the point of view ostensibly held by the writers of the functional texts. Nowhere are the methods and findings of the members of this school more radically opposed to traditional psychology than in the field of learning. They have rejected the artificial set-up of the conventional psychological laboratory, and put their greatest stress upon the immediate experiences of everyday life; they emphasize *total situations*, the *conditions* under which learning takes place, and the *uniqueness* of the individual learner.

Evidently, there is still a need for some guiding principle or frame of reference in the selection of topics for educational psychology texts.

It is the absence of such a principle, perhaps, that accounts also for the tendency of some of the functional textbook authors to include everything but the subject designated by the titles of their books. The topic, educational statistics, is a case in point. Without doubt, educational statistics is of great value to the teacher. It may be asked, however, is it a proper topic for educational psychology? In order to acquire an understanding of children, the teacher must have a knowledge of statistics, but it could be maintained, in order to master statistics, a knowledge of exponents, square root, and graphic representation is prerequisite. Then why not a chapter on algebra in an educational psychology text? On the same grounds (usefulness to the teacher), an excellent case could be made out for the inclusion of a chapter or of a whole section on the psychological novel, for, as Finney says, "one is likely to learn far more about human motives by reading Conrad, Balzac, Cabell, or De Morgan than by any technique of science."[10] Again, if intelligence-testing be included, why not also a treatment of the various psychoanalytic techniques? By what criteria is the psychology of reading a proper subject for educational psychology, and the psychology of arithmetic or the psychology of the social sciences not?

And what if all these things be treated along with the latest experimental data relative to home room organization, assembly programs, clubs, and other practical activities in which the teacher actually engages? Would the result be a *functional* book, let alone a functional

[10] Finney, R. L. "Philosophy and Science Again," *Educational Administration and Supervision*, XVI (March, 1930), p. 172.

book in educational psychology? The consensus of opinion of an increasing number of educators and psychologists seems to indicate that it would not be. There is a growing conviction that the mere piling up of facts will no more lead to intelligent application in concrete situations than did the formal, abstract verbalism against which Comenius, Pestalozzi, and Herbart long ago revolted. Consequently, more recognition is being given to the necessity for the provision of an interpretative organization of factual material—an organization that will facilitate the integration of a knowledge of facts into such a form that it can be intelligently applied. Indeed, the functional texts in arithmetic, science, and English, mentioned above, have already been criticized on the ground that they do not provide for generalization; that the material in them is too disjointed, too aimless, too irregular, and too lacking in a unifying background. They are functional only in name; they do not give the perspective that pupils need to have in their dealings with out-of-school practical problems.

So with many of the educational psychology texts called "functional." In the effort to make the findings that have accumulated from the most recent investigations in learning immediately practicable, the role of principles in application has not been sufficiently recognized. Accounts of researches connected merely by their date of publication cannot give the insight essential to an incorporation of their findings in instructional practices. The findings must be whipped into a shape such that their applicability to school learning can be tested; in other words, if they are to function in practice, the further step of relating them to fundamental principles must be taken. When this step is taken—when the products of experimentation are presented as illustrations of psychological principles, educational psychology textbooks will pass beyond the stage of "Readings in Educational Literature," "Reviews of Research," or "Bibliographies of Selected References," and become genuine functional texts.

PATTERN SECTIONED PUPILS IN LIEU OF EQUIVALENT TEST FORMS FOR CLASSROOM TESTING

HENRY F. DICKENSON
Lincoln Memorial University

Editor's note: Many persons have been interested in the problem of cheating upon examinations and cues of one sort or another. The author presents the results of an experimental study of a pattern sectioning plan for which he secured increased validity and reliability on school examinations.

AT present, due to the economic element involved, a large percent of the schools of this country are unable to furnish equivalent forms of tests and examinations. Because of the crowded conditions of the schools, there are many determiners which operate to reveal the correct response to examination items despite the teacher's efforts to the contrary. These determiners enable students to obtain clues which lead to correct answers to examination items they otherwise would not get. Students may hear the distinctive stroke of a classmate's pencil to the traditionally uniform answer type; they may get a periphery view from across the room of certain uniform movements which lead to the correct answer; they may note inflections in the teacher's voice on oral examinations; they may see the impulsive start of one who knows, as the correct answer to the multiple choice oral question is read; or, certain students may see the actual abbreviated answers which vitiate the objective results of the whole test. The last mentioned breach is evidenced by the fact that some mediocre students, if allowed to do so, tend to migrate toward excellent students for the examination period. These, and a host of others, facilitate deception and consequent reduction in reliability for the newer type objective examinations. This fact makes it imperative that sectioning within the classroom be made easier and more practical for teachers.

During the fall quarter of 1936, a member of the Lincoln Memorial University faculty asked the writer the following question: "Is there some practical way to section large classes for examination purposes without the necessity of producing equivalent forms of the examination?" No satisfactory answer was given, yet the need was apparent, and the problem soon took form. Actual classroom situations were used for experimental purposes until such sectioning technique was developed.

It was thought at the time however, that ultimately inner standards such as proper ideals would be the best approach to the problem. Confidence on the part of the teacher and its attendant effect upon the conduct of the pupils was considered. But aside from these and other con-

siderations, the fact remained that there are some students who will cheat by getting help from a neighboring paper, thereby vitiating the objectivity, reliability and general trustworthiness of the examination. In that case, only one solution seemed necessary, namely, to surround each pupil with students who were answering the same questions but in an entirely different way. Different patterns for their answers seemed a most feasible solution.

It has been the custom at Lincoln Memorial University for a number of years to administer Psychological Tests to the freshmen. The Otis-Self Administering Test of Mental Ability and the Sones-Harry High School Achievement Test furnished a basis for control and experimental groups for this study. Six hundred forty-one subjects were used in both control and experimental groups. Data from these subjects in the form of dual, triple, and quadruple choice Education tests were used in establishing the efficacy of this sectioning technique.

USING THE PATTERNED-RESPONSE SECTIONING TECHNIQUE

A class may be divided into as many sections as desired for the test period. Each section is then given a different response pattern. If four sections are desired, the teacher should have the students count off by fours, row by row, from one side of the room to the other. The ones then constitute section one, the twos section two, the threes section three, etc. Each of these sections is then presented with a different answer pattern. The type of test to be used determines the kind of pattern needed. If only a dual choice test, then each pattern needed will consist of either two letters or a two digit number. If the test is a quadruple choice type, then each pattern will consist of either four letters or a four digit number. If however, the same test should consist of three parts,—first two choice items, second three choice items and third the best of four answers,—this calls for either a four letter or a four digit number pattern for each section.

The patterns may be presented in a variety of ways. They may be presented very effectively by stamping them on the back, upside down, at the bottom of the examination, so that they may be seen by turning up the bottom of the examination when ordered to do so. The patterns may be presented by the instructor from the rear of the classroom while each section in turn faces the rear. The writer has gotten very satisfactory results from presenting number patterns by use of the fingers instead of placards. In any event it is advisable in patterning examinations to

choose patterns which have at least one different letter or number from all other patterns, rather than the same letters or numbers, so that each section may be readily distinguishable when checking the papers. Such patterns as (S-N-O-W)—(T-O-W-N)—(B-O-W-S) and (G-O-W-N) utilizes this principle. The S, T, B and G are the distinguishing letters for each section.

Now when the class has been divided into the desired number of sections and each section given a pattern, they are ready for the instructions as to how to use the patterns. Instructions should state that the first and second letters are to be substituted for true and false or yes and no or plus and minus as the case may be. If the statement appears to be true, the first letter of the pattern should be used to designate that fact. If the statement or item appears to be false, the second letter of the pattern should be used to designate that fact. As an example suppose another part of the test requires the student to pick the best of four answers. If the first answer appears to be best, then the first letter of the pattern should be used to designate that fact. If the third answer appears to be best the third letter of the pattern will designate that fact. If the fourth answer appears to be best, the fourth letter or number of the pattern will designate that fact, etc. A pre-test sample will enable the student to think in terms of his pattern rather than the traditional mode of response.

By thoughtful planning of the seating arrangement, each pupil may be surrounded by students with different response patterns.

THE EXPERIMENTAL SET-UP FOR THE SOLUTION

The experimental group (patterned response sectioned) was compared with the control group under different so called optimum conditions: (1) when the control group was taking regular oral Education examinations in the same room under regular classroom conditions; (2) when the control group was taking written Education examinations in the same room under regular classroom conditions; (3) when the control group was taking the same Education examination while in a different room; (4)[1] when control students were separated several seats apart in the auditorium while the experimental group took the same Professional Education examination under regular classroom conditions; and, finally, both experimental and control groups were rotated on the even and odd halves of an Education test under regular classroom conditions.

[1] Easley, H. Constructive Criticism offered at a session of Southern Society of Philosophy and Psychology.

That is, one group answered the odd numbered items according to one pattern, while another group answered the even numbered items according to another pattern. Then, the two groups answered the other halves of the examination without patterns. Since the control results were on the last half of the examination, the factor of fatigue had to be held constant; hence, another set-up similar to the above was used with the single exception, that the control half of the examination preceded the experimental half. The accompanying tables show the effectiveness of varying the conditions of the control group while being compared with the patterned response experimental group.

TABLE I

PATTERNED RESPONSE SECTIONED AND CONTROL TEST SCORES

	Experimental			Control			
Procedure	Subjects	Mean	S.D.	Subjects	Mean	S.D.	D/SD
Oral Ed. Test	126	36.29	3.69	126	36.94	4.00	1.33
Written Ed. Test	155	26.26	5.08	155	25.87	7.41	.54
Control Group Separated in Auditor.	126	46.02	10.74	126	46.04	10.20	.02
Rotated on Odd and Even	92	54.94	7.53	98	55.56	6.87	.59
Rotated on Odd and Even	122	82.16	9.30	122	81.00	9.36	1.19

A total of 46,194 experimental and 46,594 control responses were used in the experiment. The combined mean experimental score was 46.45,[2] with its combined standard error of .84; and the combined mean control score was 47.07, with its combined standard error of .84. This gives a critical ratio of .52 which means there is no statistically reliable difference between them, when optimum control conditions are had.

The larger critical ratios (D/SD diff.) indicate more likelihood of a higher degree of difference between the control and experimental groups. This consequently indicates a lower degree of validity (Table II r. 34), which is probably the influence of the factor of deception on the oral examinations during the first quarter.

This table shows combined validity coefficients obtained by correlating patterned response sectioned Education test scores, and control Education test scores with the several criteria. Experimental and control Education test results are correlated with criteria such as Professional Education test scores, College Grades, High School Achievement Test scores, Psychological Test scores, and High School Grades. The combined

[2] Garrett, H. E. *Statistics in Psychology and Education*, pp. 191.

TABLE II

VALIDITY COEFFICIENTS OF PATTERNED RESPONSE SECTIONED AND CONTROL TEST SCORES

	Experimental			Control		
Criteria	Subjects	r	SDr	r	SDr	D/SDr
IQ, SHA, FQG, HSG*	126	.50	.023	.34	.032	4.10
IQ, SHA, FQG, HSG	155	.47	.032	.38	.028	2.14
Prof. Ed. Test	135	.83	.027	.81	.029	.50
IQ and SH Ach.	98	.59	.068	.56	.070	.31
IQ and SH Ach.	127	.48	.067	.56	.061	.88
Combined Data**	641	.59	.021	.56	.019	1.07

* Four groups combined, and n. becomes 504.

** The square root of the mean r. square, Garrett, *op. cit.*, p. 192.

mean validity indices are r. 59 with a standard error of .021 in the former and r.56 with a standard error of .019 in the latter. The mean critical ratio for validity is 1.07. Therefore, no absolute difference is found under the conditions stated between the experimental and control groups when correlated with the several criteria. However, it is believed that the relatively larger difference between the experimental and the control of row one is due to some other factor. Here again the larger critical ratios indicate less real validity. It may be of interest to note at this point that vitiating errors of deception are indicated by the first group of oral tests in particular. Since considerable discussion, however, was provoked regarding cheating on examinations during the first quarter of the experiment, it is likely that the waning influence of this variable partially accounts for the gradual decrease in the size of critical ratios.

SUMMARY AND CONCLUSIONS

Where students sit side by side in the regular classroom, take the same examination but use different patterns for their answers, the scores are as valid, as reliable and quantitatively the equal of scores made by these same students when they sit in an auditorium several seats apart and take the examination in the traditional manner. The patterned response sectioned examinations appear consistently more valid than the types of control examinations used here. Patterned response sectioning technique, if used judiciously, will tend to eliminate deception on examinations. Such sectioning technique is feasible. It will make possible more objective and impartial testing, in the great majority of our schools, which

are financially unable to afford elaborate sectioning facilities. It will enable students to realize the instructor is dealing fairly with them by giving both sections equivalent forms, e.g., the same test. And above all else, it will discourage dependence, encourage individual achievement, increase confidence and thereby utilize more fully the nation's real wealth, the thinking ability of its citizens.

CONSTANCY OF ERROR IN LEARNING DIVISION WITH A TWO-FIGURE DIVISOR

FOSTER E. GROSSNICKLE
State Teachers College, Jersey City, New Jersey

Editor's note: Dr. Meyers, a number of years ago, suggested that errors in arithmetic were constant and not due to chance. The author offers new data and new evidence upon this important question.

AN error in division is defined as an incorrect response or a procedure which produces an incorrect result. This study was made to determine if errors in learning division with a two-figure divisor are chance errors or constant errors. Chance errors are sporadic while constant errors persist.

Studies have been made to show the constancy of errors in the fundamental operations by Myers,[1] in the basic facts by the writer,[2] and in the fundamental operations with fractions by Brueckner.[3] Except for the study by Myers, all of the other investigations cited showed that most errors were due to chance. In a few cases, however, highly specialized errors such as those resulting from the use of zero were persistent. For the reported studies, the errors were made on facts or processes that already had been learned. In this study, the errors occurred while the pupil was learning the process.

A record was kept of the errors made by 221 pupils in the second half of Grade IV. These pupils were selected from alternate classes from a group of 14 classes containing 449 pupils used in another study.[4] Each pupil was learning to divide when the divisor is a two-figure number. There were 26 practice exercises distributed over an interval of 76 teaching days. The errors reported in the following tables were made to the examples given in these practice exercises.

[1] Myers, G. C. "Persistence of Errors in Arithmetic," *Journal of Educational Research*, X (June, 1924), pp. 19-28.

[2] Grossnickle, Foster E. "Reliability of Diagnosis of Certain Types of Errors in Long Division with a One-Figure Divisor," *Journal of Experimental Education*, IV (September, 1935), pp. 7-16.

Grossnickle, Foster E., and Snyder, John H. "Constancy of Error in the Fundamental Operations in Arithmetic," *Journal of Educational Research*, XXXII (January, 1939), pp. 336-344.

[3] Brueckner, Leo J. "Persistency of Error as a Factor in Diagnosis," *Education*, XVI (November, 1935), pp. 140-44.

Brueckner, Leo J., and Ewell, Mary. "Reliability of Diagnosis of Error in Multiplication of Fractions," *Journal of Educational Research*, XXVI (November, 1932), pp. 175-85.

[4] Grossnickle, Foster E. "An Experiment with Two Methods of Estimation of the Quotient," *Elementary School Journal*, XXXVII (May, 1937), pp. 668-77.

The writer[5] previously reported a study of errors in division when the divisor is a one-figure number. In that study, the different kinds of errors were grouped under six classifications. Those classifications with a limited modification are used in this investigation. The different grouping for errors and the number of errors listed under each classification are shown in Table I.

TABLE I

CLASSIFICATION OF ERRORS AND NUMBER OF DIFFERENT ERRORS FOR EACH CLASSIFICATION

Classification	Number of different kinds
Errors resulting from estimation	2
Errors resulting from multiplication	6
Errors resulting from subtraction	7
Errors resulting from bringing down	9
Errors resulting from zero	17
Errors resulting from remainders	19
Errors resulting from faulty procedures	26
Errors resulting from lapses	27
Total	113

Table I shows that there were 113 different kinds of errors made in solving the division examples. This number does not include the various kinds of errors made in checking the examples. Many of the errors occurred so infrequently that it was necessary to eliminate them from further consideration. The writer decided arbitrarily upon either of two criteria to determine if an error should be included in this study. The error must have occurred at least 52 times, an average of two times per practice exercise, or it must have been made on six different practice exercises by one pupil. In reality only the first criterion was operative because all of the errors which satisfied the second criterion occurred at least 52 times during the 26 different practice exercises. The number of different kinds of errors which met these criteria was reduced from 113 to 24. These 24 different kinds of errors and their frequency are shown in Table II. The 26 practice exercises are divided into four parts of 6, 7, 6, and 7 exercises, respectively. This is about the same as a quartile distribution. The errors made in each of these four parts are shown in Table II.

[5] Grossnickle, Foster E. "Errors and Questionable Habits of Work in Long Division with a One-Figure Divisor," *Journal of Educational Research*, XXIX (January, 1936), pp. 355-68.

TABLE II

NUMBER AND KINDS OF DIFFERENT ERRORS WHICH OCCURRED AT LEAST 52 TIMES AND THE DISTRIBUTION OF THESE ERRORS AMONG THE 26 PRACTICE EXERCISES

Classification and Description of Error	Number of Practice Exercises 1-6	7-13	14-19	20-26	Total
I. Estimation: (Total: 1707)					
1. Estimation too small	415	219	126	139	899
2. Estimation too large	385	116	137	170	808
II. Multiplication: (Total: 2190)					
3. Wrong combination	826	351	143	109	1429
4. Carried wrong number	218	154	78	50	500
5. Did not carry when necessary	91	74	14	24	203
6. Carried when not necessary	31	13	8	6	58
III. Subtraction: (Total: 1598)					
7. Wrong combination	393	316	86	49	844
8. Forgot he had borrowed	207	197	40	19	463
9. Forgot necessary to borrow	113	93	17	9	232
10. Borrowed when not necessary	21	20	10	8	59
IV. Faulty Procedures: (Total: 764)					
11. Subtracted larger number from smaller (bottom up)	117	71	45	69	302
12. Placed first quotient figure incorrectly	153	23	15	13	204
13. Subtracted larger number from smaller (top down)	29	21	22	46	118
14. Used two-figure partial dividend when three-figure number necessary	35	13	7	16	71
15. Duplicated the quotient figure	65	4	0	0	69
V. Remainders: (Total: 696)					
16. Final remainder greater than divisor	82	72	27	46	227
17. Correct remainder in example, incorrect in quotient	57	40	12	5	114
18. Did not find final remainder	60	33	5	9	107
19. Used remainder within example greater than divisor	16	30	13	35	94
20. Final remainder is equal to divisor	47	15	12	8	82
21. Found only part of final remainder	52	9	6	5	72
VI. Lapses: (Total: 302)					
22. Used product too small, not detected in subtraction	66	22	12	10	110
23. Did not complete the example	54	34	14	4	106
24. Used correct product but wrote wrong quotient figure	66	9	9	2	86
Total	3599	1949	858	851	7257
Percent of Total	49.6	26.9	11.8	11.7	100.0

The description of each error in Table II is indicative of the error except possibly for Numbers 15 and 22. In the former, the pupil duplicated the quotient when it contained a one-figure number, as in the example $21\overline{)63}$. In that case, 33 was written for the quotient. For error Number 22, the pupil did not find the complete remainder within the example; hence he was unable to see that the estimated quotient was too small. The error is illustrated in the example on the right.

$$\begin{array}{r} 42 \\ 34\overline{)1728} \\ 136 \\ \hline 68 \\ 68 \\ \hline \end{array}$$

Table II shows that about half of the errors were made within the first six practice exercises and over three-fourths of the errors were made on the first half of the practice exercises. The persistence of error

cannot be determined from Table II but it can be shown partially in Table III.

TABLE III

DISTRIBUTION OF ERRORS PER PUPIL WHICH OCCURRED ON SIX OR MORE NON-CONSECUTIVE PRACTICE EXERCISES

Number of Error in Table II	Number of Pupils	Number of Practice Exercises			
		1-6	7-13	14-19	20-26
1	30	24	28	22	21
2	17	15	14	12	12
3	43	42	41	28	27
4	8	7	6	6	7
5	3	3	2	1	3
6	0	0	0	0	0
7	25	23	25	20	9
8	5	5	5	4	4
9	4	3	4	4	3
10	0	0	0	0	0
11	4	4	4	2	1
12	0	0	0	0	0
13	1	1	1	1	1
14	3	2	3	2	3
15	1	1	1	0	1
16	0	0	0	0	0
17	4	4	4	3	3
18	0	0	0	0	0
19	2	1	0	2	2
20	1	0	0	0	1
21	0	0	0	0	0
22	2	2	2	1	1
23	2	2	2	2	0
24	0	0	0	0	0
Total	155	139	142	110	99

Read Table: For the first error listed in Table II, there were 30 pupils who made this error on at least six practice exercises. From this number of pupils, 24 made this error on practice exercises No. 1-6, 28 on exercises No. 7-13, 22 on exercises No. 14-19, and 21 on exercises No. 20-26. Similarly, read the remainder of the table.

An error occurring on six or more non-consecutive practice exercises could not be found on any sequence of the exercises as long as at least six of them were not consecutive. Thus, if a pupil made an error to a subtraction combination on practice exercises No. 1, 2, 3, 4, 5, and 6, this error was found on six consecutive practice periods. If this same pupil, however, did not have an error resulting from a subtraction combination on any one of these six exercises, such as No. 3, but had an error of this kind on the seventh or any of the remaining nineteen practice exercises so that at least six were not consecutive, this kind of error appeared on at least six non-consecutive practice exercises. Table III shows that the errors, which occurred on at least six non-consecutive practice exercises, tended to be distributed somewhat evenly among the four parts, or

quarters, of the different practice periods. The variation between the totals for any two of the four quarters of Table III is much less than the difference in Table II between the totals for the first quarter and any of the three remaining quarters. In order to have this great variation in the totals for Table II, many of the errors must have been sporadic and not persistent.

Table III shows that there are 155 pupils who made errors on at least six different non-consecutive practice exercises. This should not be interpreted to mean that there were 155 different pupils out of the total group of 221 who made errors on at least six practical exercises. In many cases, there were several different errors per pupil as is shown in Table IV.

TABLE IV

NUMBER OF DIFFERENT KINDS OF ERRORS PER PUPIL OCCURRING ON SIX OR MORE NON-CONSECUTIVE PRACTICE EXERCISES

Kinds of Errors	Number of Pupils
0	149
1	29
2	23
3	8
4	6
5	4
6	2
Total	221

From Table IV, it is seen that there were only 72 pupils (221-149) who made errors on at least six different practice exercises. Of this number, there were 6 pupils who made 4 different kinds of errors; 4 pupils 5 different kinds of errors; and 2 pupils, 6 different kinds of errors. These 12 pupils accounted for 56 out of the group of 155 pupils in Table III who made errors on six or more practice exercises. Clearly these 12 pupils encountered great difficulty in learning division. From the total of 7257 errors in Table III, these 12 pupils made 1389 errors, or 19.1 percent of the total number. It is certain that errors made by these pupils resulted from a faulty knowledge of the number system.

The errors in Table III were made on six or more non-consecutive practice exercises. A study was made to find how many of these errors occurred in at least six consecutive practice exercises. An error would seem to be persistent if it occurred on six or more consecutive practice

periods. The results showed that one pupil made an error of underestimating the quotient on at least six consecutive periods and three pupils made combination errors in multiplication on at least six consecutive practice periods. It is difficult to be certain whether or not these multiplication errors were persistent because obviously the same combinations were not used in any six consecutive practice exercises. These data show conclusively that at most there were only four cases (very probably only one) in which an error was constant for a period of at least six practice periods out of a total of twenty-six.

Table IV shows the number of different kinds of errors per pupil when the errors occurred on at least six different practice exercises. The sporadic nature of errors can be seen from Table V. In this table, the number of different kinds of errors per pupil is shown without regard to the persistence of these errors.

TABLE V

NUMBER OF DIFFERENT KINDS OF ERRORS PER PUPIL WHICH OCCURRED DURING THE ENTIRE PERIOD OF 26 PRACTICE EXERCISES

Kinds of Different Errors	Number of Pupils
0	11
1	5
2	3
3	5
4	6
5	17
6	17
7	25
8	17
9	26
10	23
11	19
12	16
13	8
14	7
15	5
16	5
17	3
18	3
Total	221

The average number of different kinds of errors per pupil as shown in Tables V is 9.0. The average number of different kinds of errors from Table IV is 1.2. Furthermore, from Table IV it is seen that 149 pupils made no errors which occurred on at least six practice periods. Table V shows that there were only 11 pupils who made no errors. Table IV shows that only 2 pupils made as many as 6 different kinds of errors

which occurred on at least six different intervals while the average number of errors in Table V is 9. Therefore, it follows that most of the different kinds of errors made while the pupils were learning division were not persistent. Either these errors were due to faulty knowledge and the correct procedures were soon learned, hence eliminating the errors, or the errors were sporadic and due to chance. Since there were only at most four pupils who made the same error on six consecutive practice periods, the element of chance seems to have been the more potent of the two factors controlling the presence of an error.

Table I shows that 17 different kinds of errors resulted from use of zero; however, Table II shows that there was no one kind of error resulting from zero which occurred at least 52 times during the 26 practice periods. It follows that all of these different errors due to the use of zero were chance errors and not persistent. The study of errors in division when the divisor is a one-figure number showed that the errors resulting from zero were somewhat constant throughout Grades V-XV, inclusive. The percentage of error resulting from zero was 13.3.[6] In this study, the random errors resulting from the use of zero constituted only 0.3 percent of the total number of errors made. The difference between the constant and the random appearance of this type of error in the two studies was due to the method of teaching the zero difficulties. In the study when the divisor was a one-figure number, the pupils had learned division and the errors reported occurred on a test in division. Therefore, no instruction was given to the group who took the test to overcome any particular difficulty with the process. In the study when the divisor was a two-figure number, the instruction was aimed to provide for all inherent difficulties in the process. The function of zero as a place-holder was taught as well as its particular usage within an example of a given type. The very high degree of success obtained with respect to the use of zero indicates that zero should not be considered one of the formidable obstacles to pupil progress that it frequently has been portrayed to be in educational writings dealing with the teaching of number.

SUMMARY

This study shows the degree of persistence of errors made by a group of 221 pupils in Grade IV who were learning division when the divisor is a two-figure number. A record was kept of each pupil's errors on 26

[6] Grossnickle, Foster E., *op. cit.*, p. 366.

different practice periods in division. The results of this study showed the following facts:

1. There were 113 different kinds of errors made on the 26 different practice exercises but there were only 24 of these which occurred at least 52 times during the practice period. The 89 kinds of errors which did not occur an average of at least twice per exercise were assumed to be due entirely to chance, or at least they were not constant errors.

2. Out of a group of 221 pupils there were 149, or more than two-thirds of the total, who did not make the same error on six different practice exercises, but only 11 pupils made no errors in the study. These results indicate that an error was quite variable.

3. There were 72 pupils who had errors occurring on at least six non-consecutive practice exercises. Out of this group, there were 12 pupils who accounted for almost 20 percent of the total number of errors which persisted for at least six practice periods. For the errors which occurred on at least six non-consecutive practice periods, the average number of different kinds of errors for this group of 12 pupils was 4.7. Therefore, if a pupil tends to make many different kinds of errors on two consecutive practice sheets, the chances are favorable that some of these errors will persist.

4. The average number of different kinds of errors per pupil was 9.0. The average for those errors which occurred at least six times was only 1.2. When an error occurred on six practice periods out of a group of 26, it need appear in about 1 out of 4 chances.

5. The 17 different kinds of errors resulting from zero were due entirely to chance. The zero difficulties were eliminated due to the method used for instructing the pupil in the function of zero in different kinds of examples in division.

The results of this study show that the errors in division with a two-figure divisor were not persistent except possibly in a few cases for those errors which resulted from estimation and for multiplication combinations. The lack of constancy of most errors in division may be due to either of two things: first, the correct procedure is learned and supplants a faulty knowledge; or second, the error is due to chance. Since at most there were only four pupils who had an error occur in as many as six consecutive practice periods, nearly all of the errors reported in this study occurred at random throughout the period of learning division. Therefore, most errors in division are not persistent and their occurrence is largely dependent upon chance.

THE DIRECT DETERMINATION OF THE PROBABLE ERROR OF MEASUREMENT OF BINET MENTAL AGE*

GEORGE L. KREEZER
Cornell University
and
KATHERINE P. BRADWAY
Institute of the Pennsylvania Hospital

Editor's note: No individual test of intelligence has probably been as widely used as the Binet test. How reliable are the mental scores secured from this test? There have been many indirect calculations of the test's reliability. The authors supply a direct calculation of reliability.

I. INTRODUCTION

THE proper use of any measuring device requires a knowledge of its reliability—of the extent to which a given measurement may be subject to chance, accidental, or "indeterminate" errors. Investigators in various fields are therefore accustomed to specify the magnitude of the chance errors which may affect any measurement of a given property. In the physical sciences, the determination of the magnitude of chance errors can be made directly on the basis of repeated measurements of the particular property being examined. The probable error (or standard error) of a single measure, which can then be calculated, has the merit of being directly interpretable in terms of the theory of probability; it indicates the probability of the true value of a given property lying between stated limits.

In the field of psychological testing a direct procedure of repeated measurement for determining the probable error of a measurement is generally not feasible. If the same scale is administered to the same subject a number of times within a short interval, systematic errors may be introduced as a result of such effects as of practice or fatigue. If the interval of time between tests is lengthened in order to minimize effects of practice and fatigue, the factors of growth or deterioration are likely to play a role. The property measured may thus itself be altered in the course of the successive measurements, and a correct determination of the reliability of measurement is accordingly prevented.

To minimize difficulties of this sort, workers in the field of psychological testing have resorted to indirect statistical procedures to determine the reliability of a given psychological scale. If two scores can

* From the Department of Research of The Training School at Vineland, New Jersey. This study was conducted in 1937 while the authors were members of the department of research of the Training School. We are indebted to Dr. Edgar A. Doll for helpful suggestions.

be obtained on each member of a large group of subjects by means of two applications of the same scale, or by application of two equivalent forms of a scale, a coefficient of correlation (reliability coefficient) or the median discrepancy between the resulting two sets of scores may be obtained. If the data on which these statistics are based meet certain statistical requirements a probable error of measurement for the test may then be computed.[1] The obtaining of but two scores for each individual in the group reduces the difficulties that would be associated with a longer series of administrations of the same scale to the same individual.

The reliability of measurement of the Binet mental age has been determined in a number of studies by means of the indirect statistical procedures just described. A summary of the literature is given below. No direct determination of the probable error (or standard error) of measurement of Binet mental age has, however, been reported, due to difficulties of the sort mentioned in a previous paragraph.

Our specific purpose in the present study was to make a direct determination of the probable error of measurement of Binet mental age on the basis of a series of repeated measurements of the same individuals. The method we have employed to avoid the disturbing factors (as growth or deterioration on the one hand, and practice or fatigue on the other) which have previously interfered with direct determinations, has been that of using mentally deficient subjects who have reached a stable mental age level, and thus give evidence of neither growth nor deterioration in mental age. The practice or fatigue effects which may be associated with repeated examinations are then avoided or reduced through the provision of sufficiently long intervals between successive tests, without the factor of growth or deterioration invalidating the results. We have attempted in this way to determine the probable error of measurement of Binet mental age at each of a number of mental age levels.

II. SURVEY OF THE LITERATURE

A. Determination of probable error of measurement.

Determinations of the probable error of measurement of Stanford-Binet mental age (1916 scale) have been made by Otis and Knollin, by Dickson, and by Lincoln. In these studies the probable error was calculated either from the coefficient of correlation, or from the median dif-

[1] Otis, A. S. "Reliability of Spelling Scales, Involving a 'Deviation Formula' for Correlation," *School and Society*, IV (1916), pp. 676-683, 716-722, 750-756, 793-796.

ference between two sets of scores obtained by two administrations of of the scale or by the split-scale method.

Otis and Knollin[2] obtained two sets of scores for their group of subjects on the basis of the split-scale method. One hundred and eighty adults were used as subjects. The median of the set of differences between the two sets of scores was first calculated. The PE of measurement was then determined by means of the formula:

$$PE_{meas.} = \text{med. diff. divided by } 2.$$

The basic formula proposed by these authors for the determination of a probable error of measurement from the median difference between two sets of scores is:

$$PE_{meas.} = \text{med. diff. divided by } \sqrt{2}$$

If, however, the split-scale method is used for obtaining the two sets of scores, the latter formula is to be regarded as providing a PE of measurement for the half scale. To permit conversion of this value into a PE of measurement of the whole scale, the authors have derived an additional formula:

$$\text{PE of full scale} = \text{PE of half scale divided by } \sqrt{2}$$

Consequently, the formula to be used with the split-scale method becomes:

$$PE_{meas.} = \text{med. difference divided by } \sqrt{2} \cdot \sqrt{2}$$
$$= \text{med. difference divided by } 2.$$

The PE of measurement thus determined by Otis and Knollin was 5.3 months. Expressed in years this value becomes .44 years. On the basis of the number of tests available in the Stanford-Binet scale at the upper levels, the authors suggest that the true value of the probable error of an adult's score is probably somewhat greater than the value which they actually obtained; they suggest 6 months as a likely value.[3]

In a footnote (p. 121), Otis and Knollin mention a study similar to their own which was carried out somewhat later by Dickson with first-grade children, chiefly 6 to 8 years of age. The size of the group is not stated. Using the method of Otis and Knollin, Dickson found a probable error of measurement of mental age of approximately 3 months. Otis and Knollin point out that in terms of IQ points Dickson's results are comparable with their own. For an error in mental age of 3 months in a 7-year-old child corresponds to one of 3.57 points in IQ; an error in mental age of 6 months in an adult, with chronological age taken as 14 years, is likewise equivalent to an error of 3.57 points in IQ. On the basis of this correspondence, the authors consider it likely that the probable error of a mental age

[2] Otis, A. S., and Knollin, H. E. "The Reliability of the Binet Scales and of Pedagogical Scales," *Journal of Educational Research*, IV (1921), pp. 121-142.

[3] Otis and Knollin, *op. cit.*, p. 128.

score varies with the score, and that the probable error of IQ is approximately constant, being about 3.5 points for the 1916 Stanford revision of the Binet scale which they used.

Lincoln[4] obtained two sets of scores both by administering the scale twice, and by the split-scale method. He administered the scale twice and obtained two sets of scores. He then obtained two sets of scores for each administration by splitting the scale according to odd items versus even items. Finally, he obtained another dual set of scores for each administration by splitting the scale according to the first three items of a year versus the last three items of a year. He thus obtained five groups of two sets of scores. From these data he calculated five reliability coefficients, and from these coefficients he calculated five probable errors of measurement according to the formula:

$$PE_{meas.} = 0.6745 \frac{\sigma_x + \sigma_y}{2} \sqrt{1 - r}$$

in which σ_x and σ_y are the standard deviations of the two distributions of scores. The group examined consisted of 30 children between the ages of six and seven years. The interval between the examinations was 3½ to 4 hours. All reliability coefficients determined by means of the split-scale method were corrected according to the Brown-Spearman prophecy formula.[5]

The results of Lincoln's determinations of the probable error of measurement of Binet mental age were as follows: for readministration of the total scale, .11 years; for split-scale method, with first 3 tests compared with second 3 tests, .24 and .28 years; for split-scale method, with odd tests compared with even tests, .19 and .20 years.[6]

Terman and Merrill[7] have reported upon the reliability of the 1937 revision of the Stanford-Binet Scale. They obtained the average differences between IQ's on Form L and Form M of the scale for successive IQ levels. From these averages they derived the standard errors and probable errors of the IQ. From the $\sigma_{IQ.}$ they derived coefficients of reliability on the basis of the formula:

$$\sigma_{IQ} = \sigma_{dist.} \sqrt{1 - r}$$

[4] Lincoln, E. A. "Reliability of the Stanford-Binet Scale and the Constancy of Intelligence Quotients," *Journal of Educational Psychology*, XVIII (1927), 621-626.

[5] Thurstone, L. L. *The Reliability and Validity of Tests.* (Ann Arbor: Edwards Brothers, Incorporated, 1921), pp. 28 ff.

[6] Lincoln's report of probable error values is given in terms of months. We have reported them as years above, by dividing Lincoln's values by 12, in order to provide a uniform basis for comparison of results of different authors.

[7] Terman, L. M., and Merrill, M. A. *Measuring Intelligence.* (New York: Houghton-Mifflin Company, 1937), pp. 44-47.

The PE_{IQ} and the corresponding coefficients of reliability found for the successive IQ levels were: for IQ 130 and over (154 subjects), $PE_{IQ} = 3.54$ and $r = .898$; for IQ 110-129 (872 subjects), $PE_{IQ} = 3.29$ and $r = .912$; for IQ 90–109 (1291 subjects), $PE_{IQ} = 3.04$ and $r = .924$; for IQ 70–89 (477 subjects), $PE_{IQ} = 2.60$ and $r = .945$; for IQ below 70 (57 subjects), $PE_{IQ} = 1.49$, and $r = .982$.

B. Determination of coefficients of reliability.

In addition to the studies summarized above, some reports have been made of investigations in which a reliability coefficient alone was determined without the attempt being made to calculate the related probable error of measurement. Such reliability coefficients have been reported for the Stanford-Binet 1916 scale by Kelley on the basis of data supplied by Terman, and by Cuneo and Terman.

Kelley[8] reports reliability coefficients for three different groups of subjects. The indications are that the split-scale method was used in obtaining the correlations, but this fact is not clearly stated. The group of subjects in one study consisted of 108 children 8 to 9 years old. The resulting reliability coefficient was $r = .92$. The group of subjects in a second study consisted of 57 children 12 to 13 years years old. The resulting reliability coefficient was $r = .93$. The group of subjects in a third study consisted of 180 adults. The resulting reliability coefficient was $r = .93$. This third study was apparently the one carried out by Knollin which provided the data for the report by Otis and Knollin already described above.

Cuneo and Terman[9] obtained two sets of scores by administering the scale twice. The group of subjects examined consisted of 25 children between the life ages of approximately 3.7 years and 5.9 years. The interval between the two administrations of the scale was 2 days. The coefficient of correlation calculated was between the IQ scores on the two successive administrations of the scale; but inasmuch as the life ages were, in effect, the same at the two administrations, the correlation obtained can be regarded as representative of that which would have been obtained for mental age. The reliability coefficient thus determined was .95. Other correlations were obtained for sets of IQ scores obtained at greater intervals of time between administrations. We are concerned here, however, only with the question of the reliability of measurement of mental age and so shall not consider studies relating to the constancy of the IQ under conditions in which both chronological and mental age varies.

[8] Kelley, T. L. *Interpretation of Educational Measurements.* (Yonkers-on-Hudson: World Book Company, 1927), pp. 294.

[9] Cuneo, I., and Terman, L. M. "Stanford-Binet Tests of 112 Kindergarten Children and 77 Repeated Tests," *Pedagogical Seminary,* XXV (1918), pp. 414-428.

III. CONDITIONS OF THE PRESENT INVESTIGATION

A. Selection of subjects.

The experimental group used in the present study consisted of 45 mentally deficient subjects who had been in residence at The Training School at Vineland, New Jersey. These subjects were selected from a total of 600 individuals who were either living at the school at the time or had been dismissed within the previous ten years.

The following requirements were set up in the selection of subjects: (1) A subject, to be satisfactory, should show no special physical disabilities which might interfere with accurate mental age determinations—such as motor handicaps, blindness, or deafness; (2) he must be designated clinically as "no special clinical type";[10] (3) there should be at least four mental age scores determined by means of full-length Stanford-Binet examinations (1916 scale), administered at intervals not less than 1 month apart; (4) the linear mental age "growth curve" (defined below) of any subject should show that a stable mental age level had been attained, that both growth and deterioration were absent. Subjects were excluded from the experimental group if the linear growth curve showed a slope greater than + .10 or less than − .10. A slope of + .10 is equivalent to an increase of .10 of a year in mental age for each year of increase in chronological age. Similarly, a slope of − .10 is equivalent to a loss of .10 of a year in mental age for each year of increase in chronological age. To provide full scale testing of subjects with a basal year under 3 years, the Kuhlmann-Binet tests for ages under 3 years were used to supplement the Stanford-Binet scale.

Of the 600 subjects initially considered, approximately 270 had four full length Stanford-Binet mental age scores, obtained at intervals greater than a month apart. This list of subjects was then culled further on the basis of criteria (1), (2), and (4). The summary cards used for the purpose provide in condensed form, among other data, a statement of the individual's clinical type, special disabilities, and dates and scores of successive mental age examinations. Inspection of these cards permitted the list of 270 subjects to be reduced to 53 subjects through the elimination of those who were of a "special clinical type" and of those

[10] The statement that a subject is of "no special clinical type" means that he fails to show the special symptoms sometimes associated with mental deficiency and commonly used as a basis for the classification of the mentally deficient into various clinical types. Examples of such clinical types are mongolism, hydrocephalus, microcephalus, the cerebral infantile palsies, and epileptic amentia. Further details may be found in Tredgold (16).

whose successive mental age scores indicated a distinct trend upward or downward (as, for example, when each successive mental age score was higher than the preceding one.)

The final selection of subjects from the group of 53 subjects now remaining was made on the basis of the "linear growth curves" that were determined for all individuals. The term "linear growth curve" refers to the straight lines that were fitted to the set of points on a graph representing the successive mental age scores of any given subject. These successive mental age scores were plotted at intervals corresponding to changes in life age. Straight lines were fitted to the sets of points by means of the method of least squares, and the slopes of the lines thus determined by calculation.[11] The slope of the "linear growth curve" obtained for any subject thus indicates the amount by which mental age changes for a given change in chronological age (life age). The numerical value of the slope gives the change in mental age in years per year change in chronological age. It will be recalled that the slope of a line is defined mathematically as the trigonometric tangent of the angle made by the line with the x-coördinate axis. A graphic method of fitting straight lines to the sets of points was tried initially, but was given up in favor of the mathematical procedure based on the method of least squares. Eight subjects were found to have linear "growth curves" with slopes equal to or greater in absolute value than .10 and were therefore excluded from the experimental group.

TABLE I

CHARACTERISTICS OF SUBJECTS OF SUCCESSIVE MENTAL AGE GROUPS

Mental Age Group (Years)	1			2		3		4		5		6		7	
	No. of Subjects			No. of Measurements		Mental Age in Years		Initial Life Age in Years		I.Q.		Interval between Tests in Years		Slope of Linear M.A. "Growth Curve"	
	Male	Female	Total	Total	Av. per subject	Arith. Mean	Limits	Arith. Mean	Limits	Arith. Mean	Limits	Arith. Mean	Limits	Arith. Mean	Limits
0- 3	6	3	9	50	5.6	2.2	1.2- 2.8	10.2	5-23	23	9-31	1.1	.1-4.0	.033	-.09 to .08
3- 6	14	0	14	67	4.8	4.7	3.1- 5.5	16.0	7-33	35	26-44	1.4	.1-3.8	.035	-.06 to .08
6- 9	15	2	17	77	4.6	6.9	6.0- 8.1	15.4	7-42	52	46-67	1.2	.3-3.4	.051	-.05 to .09
9-12	2	3	5	23	4.6	9.8	9.0-11.5	22.4	12-37	71	64-82	1.7	.5-4.3	.048	-.01 to .08
Total	37	8	45	217	4.8	5.6	1.2-11.5	15.0	5-42	43	9-82	1.3	.1-4.3	.042	-.09 to .09

The characteristics of the final experimental group of 45 subjects are shown in Table I, with the subjects classified into 4 groups according to

[11] An exposition of this procedure is given by Guilford (10), pp. 297-301.

mental age level. For all of the subjects in this experimental group, the original Binet examination blanks were checked to insure absence of arithmetical errors in the computation of the mental age score. The successive vertical sections of the table indicate for each mental age group: (1) number of male subjects, number of female subjects, and total number of subjects; (2) total number of measures of mental age, and average number of measures for each subject; (3) mean mental age of the group and limits (lowest and highest) of the individual mean mental ages;[12] (4) mean initial life age (that is, life age at first examination) for the group, and limits of initial life ages in the group; (5) mean IQ for the group and limits of mean IQ of individual subjects;[13] (6) mean interval between successive examinations and limits of intervals for group (that is, shortest and longest intervals involved); and (7) the mean slope of the linear growth curves and the maximum negative and positive slopes for the group. The mean slopes were determined without reference to the sign of the slope of the individual curves.

B. Procedure used in calculating PE of measurement.

The $PE_{meas.}$ was determined by two methods. We may designate them as method A and method B. In *method A*, the $PE_{meas.}$ was determined for each subject individually. The mean $PE_{meas.}$ was then determined for each of the mental age groups and for the total group by obtaining the arithmetic mean of the values of the $PE_{meas.}$ obtained for the various subjects in the group. To obtain the $PE_{meas.}$ of any given subject, the following conventional formula was employed:

$$PE_{meas.} = .6745 \sqrt{\frac{\Sigma d^2}{n-1}}$$

where d equals the deviation of any mental age score from the mean mental age score of a given subject, and n equals the number of separate mental age scores obtained.

As a check on the results obtained by application of the method just described, the $PE_{meas.}$ for each group of subjects was also calculated directly from the total set of deviations found for each group (*method B*). Thus the various sets of deviations obtained for the different subjects of a group, in the application of method A, were collected together

[12] For example, subject 1 at the 0-3 year level had four Binet examinations during his residence at the institution. These were represented in the original data sheets by four Binet M.A. scores. His "mean M.A." was obtained by averaging these four M.A. scores. In this way the mean M.A. for all subjects was obtained and the mean M.A. for any one age group is the mean of these mean M.A.'s of individuals.

[13] For subjects with life ages greater than 14 years, 14 years was taken as the chronological age in the calculation of I. Q.

and the sum of the squares of these deviations obtained. The $PE_{meas.}$ corresponding to this sum was then determined by means of the formula given in the previous paragraph.

The way in which this latter procedure differs from that commonly used in obtaining a probable error or standard deviation of a distribution for a group of subjects lies in the fact that here the deviations are not obtained on the basis of a single mean score for the total distribution. Instead the deviations corresponding to the various scores of each subject are determined in relation to the mean score of that particular subject. Such a procedure permits one to determine the distribution of scores dependent on errors of measurement independently of the variations associated with individual differences among different individuals in a group.[14]

IV. RESULTS

The results of this study are summarized in Table II. The data recorded in the table were obtained through the application of method A, discussed above. The table is divided into 2 sections. In the first section (Column 2 through 5), errors are reported in terms of absolute units of years of mental age. In the second section, probable errors are expressed in terms of percentage of mental age score. The results obtained by means of method B are practically identical with those obtained by means of method A. The two sets of results are compared at the end of the following section.

TABLE II

RELIABILITY OF MEASUREMENT (IN TERMS OF $PE_{meas.}$) OF BINET MENTAL AGE IN DIFFERENT MENTAL AGE GROUPS

	$PE_{meas.}$ in absolute units (years)				In Percent (Percentage errors)[2]	
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mental Age Group (years)	Arith. Mean	σ_{Mean}	Range	Maximum Deviation[1]	Arith. Mean	Range
0- 3	.11 yrs.	.013	.04-.17	.36	5.5	2.5-12.8
3- 6	.14	.020	.04-.26	.63	3.2	.8- 6.6
6- 9	.15	.018	.03-.29	.88	2.2	.4- 3.7
9-12	.15	.011	.11-.18	.40	1.5	1.2- 2.0
Total group	.14	.010	.03-.29	.88	3.2	.4-12.8

[1] Definition of this term is given in text, page 206.
[2] Percentage error was calculated for each subject in the group individually and the results averaged to obtain means of various groups.

[14] A procedure of this sort has been used by the senior author in a number of previous studies in which it was found useful to determine the magnitude of errors of observation independently of variations in the measurements due to some factor systematically controlled. (See e.g. refs. 12 and 7).

A. Magnitude of errors in absolute units.

The chief results are shown in Column 2. This column gives the mean probable error of measurement found for the different mental age groups and for the experimental group as a whole. The mean probable error of measurement for the total group is .14 years. It is somewhat smaller at the lowest mental age level, .11 years, and increases slightly at the higher mental age levels. The differences between the mean $PE_{meas.}$ of the different mental age groups are, however, relatively small, and except for the difference between the lowest and highest mental age groups, they are not statistically significant. For the 0-3 group and the 9-12 year groups the ratio t of the difference between the means to the standard error of the difference is 2.3. According to Fisher's table of t values (6, p. 158), the difference is significant for $t = 2.179$ (for 12 degrees of freedom). The difference between the probable errors of measurement of the lowest and highest mental age groups must accordingly be regarded as significant. Column 3 of Table II shows the standard error of the mean for the arithmetical means given in Column 2.

The remaining columns of the table provide additional data which may be useful in indicating the extent to which a given mental age score may be in error. Column 4 shows the lowest and highest $PE_{meas.}$ obtained for individual subjects in the different mental age groups. Column 5 shows the "maximum deviation" found in each of the mental age groups. The "maximum deviation" represents the largest deviation found among the individuals of a given group between any mental age score and the average mental age of the subject concerned. The values given in Column 5 thus indicate the greatest amount by which a particular mental age score obtained for a given individual might differ from the "true score" (mean score) of the individual among the subjects of a particular mental age level.

As indicated above, the values obtained for the $PE_{meas.}$ in the various mental age groups are almost identical for the two methods of calculation used. Following are the results obtained for each of the mental age groups. In the case of each group, the value given first is the $PE_{meas.}$ in years, as calculated by method A; that given second is the $PE_{meas.}$ as calculated by method B: *M. A.* group 0–3, .112 years and .109 years; *M. A.* group 3–6, .144 years and .152 years; *M. A.* group 6–9 years, .153 years and .160 years; *M. A.* group 9–12, .146 years and .144 years; *Total group*, .142 years and .145 years. For all of the groups, the difference in the $PE_{meas.}$ as determined in the two ways is less than .01 years.

The empirical curve representing the distribution of the deviations (between mental age scores and the mean mental age of a given subject) for the entire experimental group is shown in Figure I. This curve represents the total set of deviations used in calculating the $PE_{meas.}$ by means of method B. The total number of deviations represented is 217. The closeness of fit of this empirically determined distribution curve to a normal probability curve was tested by means of Pearson's test for goodness of fit (10, pp. 91-93). This test resulted in a value of 7.58 for x^2, and a value of P lying between .70 and .80. This value of P indicates an exceedingly good fit between the best fitting normal probability curve and

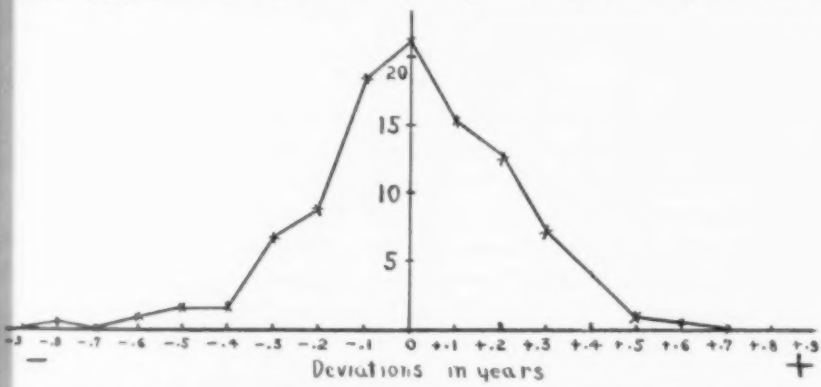

FIGURE I. Frequency polygon showing distribution of deviations of mental age scores for total experimental group. Method of determining deviations is stated in text. The points labelled along the x-axis are the mid-points of intervals extending for a distance of .05 years on each side of the point represented. Frequencies represented by the ordinates are expressed in terms of percent of the total number of observations.

the empirical distribution curve actually found, inasmuch as a statistical criterion commonly employed is that P must be larger than .05 in order for the sample to be regarded as normally distributed (6, p. 77). The value of P obtained means that there are between 70 and 80 chances in 100 that a fit as poor or worse than that obtained would be provided by a random selection of measurements such as ours, (even though the measurements were distributed normally in an infinite population). That is to say, that there are only between 20 and 30 chances in 100 that a random sample of observations would show a better fit to a normal probability curve.

B. Percentage errors.

In the last two columns (Columns 6 and 7) of Table II, the $PE_{meas.}$ is given in terms of percent. The "percentage error" was calculated sep-

arately for each subject and represents the ratio of the $PE_{meas.}$ for a given subject to his mental age, expressed in percent. Column 6 shows the mean "percentage error" for a total mental age group at the different mental age levels. Column 7 shows the smallest and largest percentage errors found in the different mental age groups.

The mean "percentage error" shows a progressive decrease from an error of 5.5 percent at the lowest mental age level to one of 1.5 percent at the maximum mental age level, with a mean percentage error for the total group of 3.2 percent. The data given in Column 7 indicate that the percentage error ranged empirically between extreme values of 0.4 percent and 12.8 percent.

V. DISCUSSION

A. The magnitude of the probable error of measurement.

The chief purpose of this study was the determination of the probable error of measurement of Binet mental age by means of a direct method of determination. The results are given in Table II and should be of value in estimating the allowance to be made for errors of measurement both in the measurement of Binet mental age of individual subjects, and in the selection for research purposes of groups of subjects intended to fall within a limited mental age range.[15]

The part of the table giving the magnitude of the probable error of measurement in terms of percent shows that the percentage error is surprisingly small, the means ranging between about 5 percent for the lowest mental age group and 2 percent for the highest, with an average value for the total group of 3.2 percent. The percentage error in measurement of Binet mental age is thus appreciably smaller than the 10 percent error regarded as characteristic of a great many measurements in biological systems.

It is of interest to compare the magnitude of the probable errors of measurement found in this study with those reported previously on the basis of the indirect methods of determination reviewed in the early part of this paper. In making such a comparison, it is helpful to distinguish between determinations which have involved the use of the split-scale method, and those dependent upon readministration of the total scale. Lincoln has reported a determination of the probable error of measure-

[15] Kreezer, G., and Bradway, K. P. "The Relation between Binet Mental Age and Motor Chronaxia," *Archives of Neurology and Psychiatry*, XXXIV (1935), pp. 1149-1171.

ment based on two administrations of the total scale to a group of children between the ages of 6 and 7. The value he obtained for the probable error of measurement was .11 years. This value is in good agreement with the value of .14 years which we obtained for our total group, and with the value of .15 years obtained for the 6–9 year mental age group, the group in this study which may be regarded as comparable in mental age to Lincoln's group of subjects.

Lincoln also treated his data by means of different forms of the split-scale procedure to obtain material for additional estimates of the probable error of measurement. This manner of treating his data provided values for the probable error of measurement appreciably greater than the value of .11 years obtained through readministration of the same scale, ranging between values of .19 years and .28 years.

The other investigators, Otis and Knollin, and Dickson, who have reported determinations of the probable error of measurement of Binet mental age, made use of the split-scale procedure alone in obtaining their parallel sets of scores. Otis and Knollin obtained a $PE_{meas.}$ of .44 years for a group of adults, and Dickson a $PE_{meas.}$ of .25 years for a group of first grade children. These values, like those of Lincoln which were based on the split-scale procedure are appreciably greater than the values for the probable error of measurement obtained in the present study. It seems not unlikely that an important factor in the smaller probable errors of measurement found in this study lies in the fact that they were based on repeated measurements with the total scale rather than on a procedure of using different parts of the same scale which are assumed to be equivalent. For if the two parts of the scale used for providing parallel sets of scores are not in fact absolutely equivalent, then the probable error of measurement which will be obtained as a result of the calculation will contain a factor due to lack of equivalence of the two parts of the scale as well as a factor due to chance errors of measurement. It is not surprising to find, therefore, that determinations of probable error of measurement based on the split-scale procedure have yielded somewhat higher values than those based on readministrations of the same scale.

B. Maximum errors involved in single mental age determinations.

One important use for probable error determinations such as are reported in the present study lies in the provision of a basis for estimating the chance error that may be involved in a single mental age de-

termination. Thus if a measurement of the Binet mental age of a given subject is made under standard conditions, a knowledge of the probable error of measurement associated with the scale permits one to state that the score obtained probably does not deviate from the correct score for the subject by an amount greater than 4 times the probable error (assuming normal distribution of scores). More precisely, the chances are 993 in 1,000 that the obtained score will not deviate from the true score by greater than that amount. Four times the probable error may thus be taken, in accordance with conventional practice, as corresponding to practical certainty that the true value will lie between the stated limits. In the present study, the value of .14 years was obtained as the average $PE_{meas.}$ on the basis of method A, and a value of .15 years on the basis of method B (after rounding to the two significant figures). We may adopt .15 years as representing the more conservative estimate; four times the $PE_{meas.}$ will accordingly equal 0.6 years. This latter value may be accepted as a suitable practical standard in estimating the size of the error to which a given mental age determination may be subject. On the assumption of a normal distribution of errors of measurement, the chances are less than 7.2 in 1,000 that a given Binet mental age will show errors of measurement greater than 0.6 years.

It is of interest to compare the estimates of the maximum error of scores made on the basis of the $PE_{meas.}$ with the maximum deviations found empirically between any given score for a particular subject and the average of a number of scores obtained by repeated examination of that subject. These maximum deviations as determined empirically are shown in Column 5 of Table II, for each of the mental age groups. If we multiply by 4 the mean probable error shown in Column 2 and compare the resultant values with the maximum deviations shown in Column 5, we find the two values are in fairly close agreement. For the lowest and highest mental age groups, the maximum empirical deviations are somewhat less than four times the probable errors of measurement; for the two intermediate mental age groups, the maximum deviations are somewhat greater than four times the probable error. Absolute agreement is, of course, not to be expected in view of the fact that theoretically a normal distribution curve will have no finite limit, that is to say, deviations of all magnitudes will be represented. The limits of an empirical distribution can be expected, therefore, only to approximate the points represented by four times the obtained probable error of a distribution. The use of the probable error determinations of the present study for

estimating the possible magnitude of errors of measurement of Binet mental ages of individual subjects or groups of subjects, presupposes, naturally, that such measurements have been obtained under the standard conditions recommended for the Stanford-Binet 1916 scale, which were the conditions involved in obtaining the data of the present study.

C. *Factors which may influence the magnitude of the probable error of measurement of Binet mental age.*

The only subdividing variable used in the present study of the precision of measurement of Binet mental age was mental age level. As shown in Column 2 of Table II, the mean probable error of measurement for successive mental age levels increases progressively with increase in mental age, suggesting a systematic influence of mental age. The calculation of a coefficient of correlation between mental age and probable error of measurement of mental age for the total group results in a Pearson r of .17.

To test the significance of this correlation coefficient, we may make use of the method recommended by Fisher for determining the significance of an obtained r.[16] For the number of degrees of freedom represented by our data, r to be significant should be equal to at least .288. The obtained r of .17 found by us for the relation between M.A. score and the probable error of measurement is appreciably smaller than .288 and therefore must be regarded as not significant. More precisely, the chances are greater than 5 in 100 that a correlation as large as .17 could arise by random sampling from an uncorrelated population.

Other conditions besides mental age level which might be considered for their effect on the probable error of measurement of Binet mental age are chronological age and sex. Within the limits of variation of these properties in the present study, neither of them was found to have any influence on the precision of measurement of mental age. Thus, the coefficient of correlation between chronological age and $PE_{meas.}$ of mental age, after correction for variations in mental age was .03, indicating the absence of any significant relationship. Comparison of the probable error of Binet M. A. for the two sexes showed no appreciable differences between them. In the total experimental group, the mean $PE_{meas.}$ was approximately the same for both males and females, .14 years for males and .16 years for females.

Apart from the possible factors mentioned above, the variability of measurement of Binet mental age scores must be regarded as a function

[16] For an exposition of this procedure, see Fisher (6), 159 f.

of the resultant effect of the host of factors that must be regarded as operative both in the subjects examined, and in the procedure of administration of the Binet scale. Inasmuch as the data used in the present study were obtained over a period of 10 years, and represents the participation of a considerable number of different examiners, the variability in results that might be produced by inevitable individual differences in the method of administration must be regarded as one of the factors present.

Finally, the question may be raised of the extent to which the determinations of $PE_{meas.}$ made in this study were influenced by changes in M. A. due to growth or deterioration. It will be recalled that subjects who met the other criteria were included in the experimental group if the slopes of their linear "growth curves" fell between the limits of $+.10$ and $-.10$. It was thus arbitrarily assumed that slopes falling within this range represented subjects in whom M. A. change was substantially absent. If one regards such small positive or negative slopes as representing real changes in mental age, then it is necessary to conclude that the true probable errors of measurement of Binet mental age are even somewhat smaller than those determined. For in the calculation of deviations of different M. A. scores from an average value for a given subject, no attempt was made to separate a component due to chance errors and a component due to true growth or deterioration. In accepting the determinations of the present study, one is thus erring if at all, on the side of safety.

VI. SUMMARY

The purpose of the present study was to determine the probable error of measurement of Binet mental age by means of a direct method of calculation. The study is regarded as of interest not only in connection with the specific results obtained, but as illustrating a method which can be used to supplement the indirect methods (based on calculation of coefficients of correlation and median differences) by means of which the probable errors of measurement of mental age scales have previously been determined. A direct method of determination of the error of measurement of any property requires repeated measurement of the particular property being examined. This procedure has not been previously considered generally feasible in determining the precision of measurement of mental age scales since (except for adult individuals) the subjects examined would be expected to change with growth, or if repetitions

were close together, the results would be likely to be modified by practice or fatigue effects. The method adopted in the present study for avoiding these difficulties consisted in the use of mentally deficient subjects who had reached a stable mental age level. Repeated examinations could therefore be made at intervals far enough apart to avoid practice effects, and yet avoid difficulties due to growth or deterioration.

The experimental group used consisted of 45 subjects ranging in mental age from 1 to 12 years. These subjects were selected from an initial group of 600 individuals. An average of 4.8 mental age scores (1916 Stanford-Binet) was obtained for each individual in the group, and the probable errors of measurement calculated for the individual subjects. The average probable error of measurement for the total experimental group was found to be .15 years. The magnitude of the $PE_{meas.}$ showed no significant correlation with mental age level, although a significant difference in the probable error of measurement between the lowest and highest mental age levels was found. The values of the probable errors of measurement of Binet mental age obtained in this study were found to be smaller than the values obtained by previous investigators by means of indirect methods involving the use of the split-scale procedure, but agreed closely with a previous determination by Lincoln involving the readministration of the total scale. The value of 0.15 years obtained for the $PE_{meas.}$ indicates that the chance of an individual Binet mental age score being in error by an amount greater than 0.6 years (4 times the $PE_{meas.}$) is very slight, less, in fact, than 7.2 chances in 1,000, if a normal distribution of errors is assumed. A value of 0.6 years thus provides a convenient value for use in practice in estimating the magnitude of the error that may be involved in a given Binet mental age score.

BIBLIOGRAPHY*

1. Berry, C. S. "A Comparison of the Binet Tests of 1908 and 1911," *Journal of Educational Psychology,* III (1912), pp. 444-451.
2. Bobertag, O. "Uber Intelligenzprüfungen," *Zsch. f. Angew. Psychol.,* VI (1912), pp. 495-538.
3. Cushman, C. L. "A Study of the Reliability of Mental Tests as Used in Oklahoma City," *Journal of Applied Psychology,* XI (1927), pp. 509-511.
4. Doll, E. A. *The Growth of Intelligence.* (Princeton: Psychological Review Company, 1920), pp. 130.

* References to previous papers which report, specifically, upon reliability of Binet mental age are given in footnotes 2 to 9, and discussed in section II.

5. Durling, D. "Note on the Comparative Reliability of the Stanford-Binet below the Age of Six," *Journal of Applied Psychology*, XVI (1932), pp. 331-333.
6. Fisher, R. A. *Statistical Methods for Research Workers.* (London: Oliver and Boyd, 1928), pp. 269. (Second edition).
7. Glanville, A. D., and Kreezer, G. "The Maximum Amplitude and Velocity of Joint Movements in Normal Male Human Adults," *Human Biology*, IX (1937), p. 205.
8. Goddard, H. "Reliability of Binet-Simon Measuring Scale of Intelligence," from *Transactions of the Fourth International Congress on School Hygiene*, Buffalo, V (1913), pp. 693-699.
9. Goodenough, F. "A Critical Note on the Uses of the Term 'Reliability' in Mental Measurement," *Journal of Educational Psychology*, XXVII (1936), pp. 173-178.
10. Guilford, J. P. *Psychometric Methods.* (New York: McGraw-Hill Book Company, 1936), p. 566.
11. Kelley, T. L. "The Reliability of Test Scores," *Journal of Educational Research*, III (1921), pp. 370-379.
12. Kreezer, G., and Glanville, A. D. "A Method for the Quantitative Analysis of Human Gait," *Jour. Genet. Psych.*, L (1937), pp. 127-129.
13. Madsen, I. N. "Reliability and Validity of Stanford-Binet Tests when Administered by Student Examiners," *Journal of Educational Research*, XXVIII (1934), pp. 265-270.
14. Otis, A. S. *Statistical Method in Educational Measurement.* (Yonkers-on-Hudson: World Book Company, 1925), pp. 247-256.
15. Rugg, H. O. "Statistical Methods Applied to Educational Testing." *Twenty-First Yearbook, National Society for the Study of Education.* (Bloomington: Public School Publishing Company, 1922), pp. 45-91.
16. Tredgold, A. F. *Mental Deficiency.* 5th edition. (New York: William Wood and Company, 1929), pp. 204-292.

EDITORIAL

AN ETHICAL RESPONSIBILITY OF THE ABSTRACTOR

THE very quantity of educational research defies even the most systematic reader to read even any considerable proportion of it. The majority of it is unpublished. Much unpublished research is probably more valid than much that attains printers ink. One can hope to keep abreast of that which is of most concern to him only (1) by obtaining as interlibrary loans unpublished theses listed in the annual compilation published by the United States Office of Education and (2) by reading available abstracts which appear in the *Loyola Educational Digest*, the issues of *Review of Educational Research*, in the *Journal of Educational Research*, and elsewhere.

Unfortunately, those who prepare reviews, abstracts or summaries of educational research do not sometimes discriminate carefully between investigations the conclusions of which may be relied upon and those which by reason of the nature of the basic data on the method of treatment, or both, yield conclusions which may be far from truth. For example, the replies to a questionnaire may be from respondents who are decidedly unrepresentative of the entire group of which they are supposed to be a good sample; differences between two methods of teaching may be the result of factors not controlled in the experiment or the differences may be only in respect to certain easily measurable types of educational outcomes; the numbers involved may be so small as to make conclusions statistically reliable. For one or more of these reasons no differences may be reported between two methods of teaching, large and small classes, and other experimental variables when real differences actually exist.

In preparing reviews of research in mathematics for issues of the *Review of Educational Research*, the author has seen fit to discard more than half of the studies available. To have reported the conclusions of these studies would have served to mislead teachers and administrators.

A much greater responsibility for critical selection rests upon the reviewer or summarizer of research than upon the editor of a journal in which fairly complete reports of studies are reported by investigators.

In the review, much more often than not, the details of the investigation which reveal its flaws and weaknesses are not given. The reader has little or no opportunity to judge the reliability or the validity of the basic data. He must accept the conclusions on faith. That faith should be well founded.

With research as abundant as it is in education in these years, there is great need for reviews and summaries of research. With the great mass of field workers so ill prepared to evaluate critically the data, methods and conclusions of educational research as are teachers, the need is multiplied. There is for the same reasons an unusual responsibility for critical selection and interpretation by the reviewer. This is accentuated by the fact that because of artificial pressures upon teachers, administrators, graduate students and college teachers of education, there is an enormous amount of research being conducted upon educational problems which will not withstand careful scrutiny.

There is ample evidence that reviewers in whom teachers and other students place much confidence are neither selective nor critical and are perpetuating and spreading conclusions, which if true are not established by the investigations from which they are derived. The conclusions of studies which upon careful examination prove to be clearly unreliable may be found in reviews published by men and organizations from whom we have a right to expect greater scholarship. This situation should be faced squarely and education protected from a practice which can serve only to spread error and confusion and to bring about deserved discredit from scholars in other fields.

HARL R. DOUGLASS.

REVIEWS

Address all communications relative to book reviews to Editor Kai Jensen, University of Wisconsin, Madison, Wisconsin.

CURTI, MARGARET W. *Child Psychology* (New York: Longmans, Green & Co., 2nd. ed. rev., 1938), pp. vii + 458. $3.20.

The unique merit of Curti's new revised child psychology arises from the fact that it advances no claims to eclecticism, the favorite sales slogan of the day, but instead vigorously champions the position that the meaning or significance of behavior items is determined by the point of view which is held. Thus the data of child psychology receive varying interpretations according to the particular school of psychology and, we may add, the fundamental theories of the different schools are frequently tested by the data of child development.

The author briefly reviews and criticizes the various psychologies under functionalistic and structuralistic categories, accepting what is termed monistic functionalism as the integrating point of view in her work. Adherents of other viewpoints in psychology will probably dissent from some of the evaluations made and not all psychologists will admit that monistic functionalism and organismic psychology are one and the same. Indeed the Gestalters and some behaviorists advance similar claims to an organismic approach. Further, in all fairness to so-called structural psychologies, we question the validity of the following as a criterion distinguishing structuralism from functionalism:

"Structuralism tends in general to be descriptive and more or less static; functionalism seeks rather to explain in the sense of discovering dynamic relationships" (p. 22).

Also the author, although commenting upon the almost inescapable mind-body dualism of our culture, defines psychology in the following passages now with an eye to monism and again in terms implying dualism:

"The subject-matter of psychology is neither consciousness nor behavior (in the narrow sense), but the adaptive activity of the whole human being—activity which may or may not be conscious, but which is always unified, that is, activity of an organism" (p. 28).

"Child psychology as a science involves primarily the attempt to gain a useful and hence analytical understanding of the chief problems of mind and its development through childhood" (p. 32).

The dilemma of those who inveigh against analysis or attention to parts and then are stopped short when they attempt to advance by contemplation of wholes is illustrated by the following:

"A functionalistic approach differs from a structuralistic approach first in emphasizing, if not in so many words then by implication, the fundamental importance of 'wholes' rather than elements; and second in treating the various phenomena in the field in relation to other past and present phenomena which are also parts or aspects of the whole" (p. 22).

"although it is the whole child that we want to understand, we cannot get such understanding by studying the whole child all at once and all the time. For the sake of convenience and effectiveness

in reaching serviceable solutions of the problems in the field it is necessary to deal with the material under separate topics and to isolate particular phases of larger problems for special study'' (p. 33).

From the preceding discussion it is evident that the introductory chapter is a sophisticated approach to the study of child development appropriate for the serious and mature student able to realize that fundamentally ''the most important aim of child psychology is simply to understand mental development in the child'' and to comprehend that ''One of the canons of scientific method, namely that scientists should not as scientists attempt to set standards of value, is frequently violated by writers on child psychology when they tell parents and teachers how children *should* be brought up, and thus lead these grown-ups to believe that they are getting scientific guidance, when as a matter of fact they are simply being given the personal opinions of the psychologist. Now the type of control that is actually exercised will depend entirely upon the values in human life which are considered worthy of attainment'' (p. 4).

In contrast with the introductory chapter is the naturalistic diary-type account of Ann's first five years. Pedagogically considered, this is an interesting device because it gives the beginning student a running description of the development of a particular child preparatory to a more complete and more abstract analysis of human development. The reviewer is, however, somewhat disturbed to find *interpretations* in this material which are introduced as those of the child's ''psychological mother'' and thus bear, for the non-discriminating reader, the stamp of sophistication. As instance of this we have:

''Is this smiling little animal already thinking? Her turning toward the stairs and waiting might be a mere set of conditioned responses—but it looks like the dawn of ideas'' (p. 36).

''For some time grown-ups had played peek-a-boo *for* Ann, managing the cloth themselves. Now she plays the game herself. Surely this means *thought* on the part of the baby. Her offering of food to her friend, and her putting of the cereal spoon in her *mother's* mouth, are corroborating incidents'' (p. 39).

A second major contribution, provocative and stimulating to the serious student of child psychology, is the author's discussion of intelligence. What is the nature of intelligence? What effective relation does it have to environmental circumstance? How does it alter with increasing age? These are all live issues for psychology and education and Curti's treatment should lead to caution in the ready acceptance of easy generalizations.

The author is not consistent in precisely defining and adhering to a definition of intelligence and intelligence tests. On the one hand we note that the score on intelligence tests rests in part, at least, upon what has already been learned; that in essence these tests are but indirect measurements based upon the assumption of a common effective environment shared by all taking the tests. Thus ''this description of the intelligence tests most commonly used makes it clear at once that the IQ is not a direct measure of inherited intelligence, but that it depends in part upon acquired knowledge, and particularly upon ability to handle language'' (p. 167). Mental testers will of course add that the IQ is not in itself a measure but merely a ratio.

Following this the author reviews the results of tests administered to identical twins reared apart and implies that the differences in score result from environmental modifications of intelligence.

"Had some of these separated twins been brought up in a remote hamlet of East Tennessee, or in certain orphanages, the differences both in intelligence and social adaptability would almost certainly have been more striking" (p. 175).

"The fact that as large differences in intelligence as this exist, when the environment of the separated twins were not greatly different, affords unquestionable evidence that environmental conditions operating over a period of years may definitely affect the IQ" (p. 177).

And only on page 178 does the author clearly state that "hereafter when we say intelligence we shall always mean intelligence test score." From this there seems to be a further step to the assumption that such "intelligence" represents "the actual degree of adaptability of the individual." Thus:

"If it be argued that after all intelligence—actual innate capacity as determined by the genes—is not changed at all, we will agree, granting that definition of intelligence. But what society is interested in after all, and what parents and children are interested in, is not a hypothetical capacity that once may have existed, but the actual degree to which the individual concerned can adapt himself successfully to the complex conditions of living in this world" (p. 182).

But if the environment is indeed so potent in altering the degree of intelligence of an individual why so quickly dash the hopes of those who would call for an effective realization of such results?

"On the other hand parents should remember what all teachers know, that there is a wide distribution of natural intelligence, and that hence most children cannot be expected to improve their IQ's significantly" (p. 195).

"Mental age, representing the limit of attainment in a number of abilities, does not increase after maturity" (p. 195).

"In case no special handicap exists, and there is no improvement in IQ over a lapse of years even under the most favorable conditions, parents will then realize that their child is simply dull, or average, or merely fairly bright" (p. 196).

Likewise challenging to the generally accepted views regarding the nature of growth curves is the author's assertion "that mental growth tends to be linear, to go on at the same rate throughout childhood" (p. 188).

In general the work places relatively greater stress upon the so-called higher mental processes, thought and reasoning, etc., than upon the development of sensory-motor processes. Curti is aware that many will not find it possible to go all the way with Piaget in his interpretation of reasoning in children but she has presented an admirable exposition of his views and one which should cause students to see that these aspects of development deserve just as careful study as do sequences of locomotion, manipulation, and so on.

In conclusion the reviewer would call the attention of educators to the fact that the doctrine of "felt needs" finds scant support among child psychologists. As Curti expresses it:

"Neither can we assume that if the organic needs are adequately met, the child will develop rich and abiding motives of other sorts. There are ways of dulling interests, there are means of restricting the development of the child, which are employed, though not with that intent, in the best families and in the best schools and in the most civilized countries" (p. 132).

KARL C. PRATT
Central State Teachers College
Mt. Pleasant, Michigan

JENSEN, ARNE S. *Psychology of Child Behavior* (New York: Prentice-Hall, Inc., 1938), pp. xxi + 664. $3.85.

"To
my students in child psychology
in appreciation of
their splendid spirit of coöperation
in the production of this book"

In this dedication the author, an associate professor of psychology in a nor

mal school, makes unwitting comment on the characteristics and limitations of his text. The writing of a child psychology as a coöperative enterprise in which typical normal school students participate, in the selection and inclusion of materials or serve as guinea pigs to determine what may be comprehended pleasantly and painlessly, is foredoomed to superficiality, triviality, misleading simplification and a lack of critical discrimination in selection and organization. The participation of such a group will produce a sentimental and fuzzy inspirational tone rather than a clear-cut descriptive approach. Controversial issues which might lead to disturbing doubt will not be permitted to intrude and corrode the pure serene wonder of contemplation of things-in-general, which to anyone but a "school-marm" obtaining more credits in a summer term at a normal school would be obvious and hackneyed.

As one might anticipate, the work scarcely ever gets down to the bed rock of data obtained from experimental researches into the nature and characteristics of child behavior and development but instead rests upon a second or third hand or upon acquaintance even more removed. It is a compilation of the generalizations and generalities of other compilers rather than a reinspection of original data. Insouciant as to latitude the book wanders like a Swiss Family Robinson through the fields of general or introductory psychology, social and educational psychology, sociology, mental hygiene, child guidance, et cetera.

For the author, this compilation or alleged integration knowing no boundaries or compartments, so much the vogue in "progressive" circles, is eclecticism. Thus

"the approach throughout has been rather eclectic. No school of psychology has been followed; no material has been left out because it came from a particular school" (p. vii).

"Some of the different schools of psychology differ more in terminology than in actual treatment of subject matter. The eclectic method may overlook such verbal differences and fuse into one whole the actual facts gleaned by a variety of researchers" (p. 542).

It is apparent that Jensen believes that schools of psychology differ in verifiable data and in terminology. For him electicism consists in collecting all data, employing a common terminology, and organizing the materials into a whole. The reviewer holds that the collection and organization of materials or data is not eclecticism (see Warren's Dictionary). Organization of data proceeds from a frame of reference. Consistency in terms of this reference provides an integration, a school or emphasis of view. Most alleged eclectic treatments are actually syncretistic. Our author is no exception—he will have his cake and eat it too, as is shown in his position in regard to race, instinct and emotion.

For him the focus of attention is upon what he terms "behavior hygiene" rather than upon a systematic and descriptive child psychology. For

"As long as such results (referring to scientific of pure studies) are merely tabulated, not actually made a part of the tools of scientific procedure in dealing with the problems of childhood, they are rather negative—at least not positive so far as the welfare of the child is concerned" (p. 5).

"the scientific procedure in child guidance — behavior hygiene — must revolve around and continually come back to the child for further study and evaluation" (p. 10).

"It is then as a positive, active, directing agency in child development that child psychology makes its demand for recognition as a factor in producing a vigorous, healthy childhood, which in turn

will guarantee a national vitality of unbounded energy and possibilities'' (p. 9). Elsewhere, somewhat ambiguously, immediate aims are distinguished from ultimate aims:

''The immediate aim of this study is to learn to understand the child; another aim, more remote, is to form the basis upon which later development is built'' (p. 40).

In spite of verbal disclaimers the systematic position is that of interactionary dualism:

''The child has been considered as a unique individual, as a whole person. The term *biomental unit* has been applied to avoid the inference of dualism'' (p. vii).

''The field of child psychology is the investigation and understanding of the process of development of mental life and behavior of the child'' (p. 39).

''The tensions due to birth trauma, hunger, elimination problems, and disturbances of sleep periods, are mainly psychic. They affect the health of the body as well as that of the mind'' (p. 546).

''The general learning process almost inevitably leaves a definite stamp on the psychic part of the child'' (p. 546).

''The ideational and idealistic elements are largely psychic with strong emotional elements attached'' (p. 547).

''The stimulus comes to the child; the impulse reaching the brain causes the sensation; now the correlation centers of the brain seem to receive this new sensation, mull it over carefully, compare and weigh it against previous experiences and ideas, and then finally issue an order to the motor equipment to respond in some manner. The suitableness of the response is determined not only by the immediate sensory stimulation, but more definitely by the relationship discovered between it and previously stored ideas'' (p. 291).

Errors of fact or statement provide the following curiosities:

''We are dealing not only with living organisms in this field, but perhaps with the most difficult of all forms of life'' (p. 41).

''It (psycho-physical method) resembles the testing method in that it compares the various phases of growth of the child and attempts to discover relationships between them'' (p. 49).

''The various organs of the body are developed during the period of the fetus . . . '' (p. 61).

''At birth, girls are slightly taller and heavier than boys. . . .'' (p. 65).

''During the first week after birth the rate of growth in motor activity and organization perhaps exceeds that of any other week in child's life'' (p. 78).

''The tactile, gustatory, olfactory, and auditory senses, as well as the less definitely localized senses of equilibrium, temperature and others, are ready to function in an elementary manner in a few days after birth'' (p. 152).

''The infants born into the world are all different from one another. Each one is a unique being differing in every measurable trait from all other individuals'' (p. 173).

The purpose of the school is to socialize the child and socialization is conformity of thought:

''The whole school spirit should be imbued with the ideal of producing a strong, self-reliant, capable, well-balanced generation whose behavior patterns are thoroughly socialized'' (p. 600).

''In some cases the person with superior mental and physical equipment becomes a menace to society. In order to become socialized the child must master the life and culture of the group and must learn willingly to conform; . . . Social adaptation is much facilitated by using common speech. This is in a deeper sense than merely vocabulary; it means that expressions are in conformity with the customary trend of thought'' (p. 589).

The status of the child in society is viewed through Pollyanna spectacles which not only provide dewy rose tints but induce considerable astigmatism.

''During the long, dim centuries when prehistoric man was just emerging from the mists of the past, the child appears to have received little respect and to have been rated as of little value. As a

general rule the more primitive condition of the people, the less consideration did the child receive. Wherever the group attained some degree of culture, the condition of the child tended to improve; nevertheless, the lot of the child in early human society was not an enviable one'' (p. 18).

''Perhaps the darkest picture of the child in history must be drawn from the child's enslavement in the industrial world'' (p. 23).

''Out of these ages of dark ignorance, of abuse, of domination by a superstitious, cruel, grasping society, the child has finally emerged'' (p. 23).

''the child could no longer be ignored, and he was rapidly moving toward the center of the modern arena'' (p. 34).

We are assured repeatedly and somewhat unnecessarily that the child is a precious possession:

''We are actually beginning to realize that the child is the most precious possession of the race'' (p. 10).

''it laid the foundation for the modern conception of the child as the most precious possession of the race'' (p. 32).

''make life worthwhile for children, the most precious possession of the race'' (p. 34).

''As the great value of the child was emphasized in a previous chapter, it is not necessary to repeat that the child is worth all the attention we can bestow upon him'' (p. 40).

''thinking people are coming to the realization that the most valuable thing in all the world is the child'' (p. 196).

''with the general acceptance by society of the child as the most precious possession of the race'' (p. 622).

Lastly, we are informed that there are no limits to man's possibilities.

''By virtue of this creative activity on the part of the human organism, a position exalted above any other creature is attained. Activities beyond the most fanciful imagination of historic man are possible. Instead of moving along paths and patterns laid down in the nervous organism and having definite limits, we may conceive of man's responses as being as limitless in scope as the entire universe'' (p. 153).

KARL C. PRATT
Central State Teachers College
Mount Pleasant, Michigan

RESEARCH ABSTRACTS and BIBLIOGRAPHIES

Address all communications relative to research abstracts and bibliographies to A. S. Barr, University of Wisconsin, Madison, Wisconsin.

BULLETINS

ACKLEY, CLARENCE E. *Teacher Certification of Instructional and Administrative Employes in the Public Schools of Pennsylvania, 1937–1938*, Research Service in Education, Bulletin 74, Research Circular No. 9 (Harrisburg: Department of Public Instruction, Commonwealth of Pennsylvania, 1939), 28 pp.

This bulletin presents a detailed analysis of teacher certification facts in the State of Pennsylvania.

ADAMS, THURSTON. *Motion Pictures in Physical Education* (New York: Bureau of Publications, Teachers College, Columbia University, 1939), 57 pp.

This bulletin discusses the place of the motion picture in physical education, producing and using the experimental films, equipment for taking and projecting motion pictures, and suggestions for using motion pictures.

EDUCATIONAL RESEARCH SERVICE. *Questionnaire Studies Completed—No. 10, 1938-39*, Circular No. 8 (Washington, D.C.: American Association of School Administrators and Research Division of National Education Association, September, 1939), 11 pp.

This circular presents a digest and bibliography of 59 questionnaire studies submitted to city and state superintendents of schools.

GATES, A. I., BOND, G. L., and RUSSELL, D. H. *Methods of Determining Reading Readiness* (New York: Bureau of Publications, Teachers College, Columbia University, 1939), 55 pp.

The authors present the results of a statistical study of the predictive value of certain measures of reading readiness.

LATON, ANITA D., and BAILEY, EDNA W. Assisted by JOSEPH SCHWAB. *Suggestions for Teaching Selected Material from the Field of Genetics*, Monograph No. 1, Bureau of Educational Research in Science, Columbia University (New York: Bureau of Publications, Teachers College, Columbia University, 1939), 66 pp.

This monograph contains a summary of scientific materials relative to the subject of genetics prepared for elementary and secondary school teachers.

LUNN, J. E., et al. *Comparison of Pupils Who Participate in No Extra-Curricular Activities with Those Who Take Part in Three or More Such Activities*, Pupil Personnel Study of Pupils in Minnesota Public Schools, Part I, Section 7 (St. Paul: Minnesota State Department of Education, 1939), 12 pp.

The authors contrast pupils who participate in no extra-curricular activities with those who engage in three or more activities in intelligence quotients, grade placements, incidence of problem cases, quality of work, and personality traits.

LUNN, J. E., et al. *Relation of Grade Placement to Other Factors in Education,* Pupil Personnel Study of Pupils in Minnesota Public Schools, Part I, Section 8 (St. Paul: Minnesota State Department of Education, 1939), 27 pp.

The authors compare mentally retarded and mentally accelerated pupils in various functions.

LUNN, J. E., et al. *Relation of Quality of Work to Other Educational Conditions,* Pupil Personnel Study of Pupils in Minnesota Public Schools, Part I, Section 9 (St. Paul: Minnesota State Department of Education, 1939), 19 pp.

The authors trace the influence of the quality of school work in retarded and accelerated pupils.

LUNN, J. E., et al. *Comparison of Good and Poor Achievers as Problem Cases,* Pupil Personnel Study of Pupils in Minnesota Public Schools, Part I, Section 10 (St. Paul: Minnesota State Department of Education, 1939), 26 pp.

Poor achievers as problem cases are contrasted with the best achievers.

MCPHERSON, ORPHA. *Summer Vacation Activities of One Hundred Farm Boys and Girls in a Selected Area* (New York: Bureau of Publications, Teachers College, Columbia University, 1939), 74 pp.

After describing the procedure, the author discusses conditioning factors, responsible work, reading, music and radio, play activities and individual interest, nature experiences, social contacts, and provision made by the school.

MEECE, LEONARD E., and SEAY, MAURICE F. *Financing Public Elementary and Secondary Education in Kentucky,* Bulletin of the Bureau of School Service, Vol. XII, No. 1 (Lexington, Kentucky: University of Kentucky, September, 1939), 182 pp.

The authors present data relative to the financing of elementary and secondary schools in Kentucky for the use of educators and legislators.

PALMER, EMILY G. *Occupational Trends in California with Implications for Vocational Education: VI. Trends in Service Occupations — Personal, Business, and Repair,* Bulletin of the California State Department of Education, No. 7 (Sacramento: California State Department of Education, July, 1939), 138 pp.

This bulletin summarizes occupational trends in the service occupations: personal, business, and repair services.

SANDS, LESTER B., and ALMACK, JOHN C. *History of Education Chart* (Stanford University, California: Stanford University Press, 1939), 14 pp.

This chart is an outline for students pursuing a first course in the general history of education and for graduates who are looking for direction in reviews. Topics are arranged in chronological order to furnish a picture of movements and national developments.

SCHWEHN, HILDA, REMMERS, H. H., and WHISLER, L. D. *Guidance in the Secondary Schools of Indiana,* The Division of Educational Reference, Studies in Higher Education, XXXVI (Lafayette, Indiana: Purdue University, June, 1939), 28 pp.

The authors discuss various measurements made in the schools, persons administering and using various tests, records kept and percentage of use by staff persons obtaining various records, percentages of various organizations in 196 small and 139 large schools, various

types of health guidance given and staff members participating, frequency with which various curricula and courses are offered, educational and vocational guidance, various guidance functions carried on and staff participation in guidance functions in schools with guidance programs.

SMITH, HENRY LESTER, and EATON, MERRIL T. *Twenty-Sixth Annual Conference on Educational Measurements*, Bulletin of the School of Education, Vol. XV, No. 4 (Bloomington, Indiana: Bureau of Coöperative Research, Indiana University, September, 1939), 53 pp.

This report contains a discussion of measurement in music, present tendencies in educational measurement, measurement of speech and hearing defects, reliability of the categorized and uncategorized mental test, some assumptions underlying the educational measurement movement.

WAGNER, JONAS E., CAMPBELL, DONALD W., and MORNEWECK, CARL D. *Salaries of School Employes 1937–1938*, Research Service in Education, Bulletin 74, Research Circular No. 11 (Harrisburg: Department of Public Instruction, Commonwealth of Pennsylvania, 1939), 27 pp.

The authors present an analysis of salary conditions in Pennsylvania.

WILEY, LESTER E., and WILEY, A. McBROOM. *Measurement of Educational Progress*, Bulletin of Research Activities, R-3 (Columbus, Ohio: State Department of Education, September, 1939), 80 pp.

The authors present data relative to the validation and use of instruments for the measurement of pupil growth by means of the Ohio Every Pupil Tests.

WOODWARD, ANNIE C., et al. *Organizations Interested in International Relations*, Research Bulletin, Vol. XVII, No. 4 (Washington, D. C.: National Education Association, September, 1939), 54 pp.

This bulletin contains the name, purpose, and other pertinent facts relative to organizations interested in international relations.

NEW TESTS

BRUNER, HERBERT B., LINDEN, ARTHUR V., and WOOD, HUGO B. *A Tentative Check List for Determining Attitudes on Fifty Crucial Social, Economic, and Political Problems* (New York: Bureau of Publications, Teachers College, Columbia University, 1939), 15 pp.

The Check List contains two hundred and fifty statements on fifty social, economic, and political problems. Each statement is intended to approximate a "reactionary," "conservative," "Middle-ground," "progressive," or "radical" position. The test is for use by college students, teachers, and adults.

ENGELHART, MAX D. *June and August, 1939 Comprehensive Examinations and the September, 1939 Placement and Exemption Tests of the Chicago Junior Colleges* (Chicago: Department of Examinations, Chicago City Junior Colleges, 1939), English Qualifying Examination, 19 pp. Biological Science 101-102, Essay Section, 7 pp., Biological Science 101-102, Objective Section, 16 pp., Humanities 201-202, Essay Section, 8 pp., Humanities 201-202, Objective Section, 19 pp., Physical Science 101-102, Essay Section, 7 pp., Physical Science 101-102, Objective Section, 14 pp., Social Science 101-102, Essay Section, 11 pp., Social Science 101-102, Objective Section, 19 pp., English

Placement Test, Form C, 8 pp., Biological Science, Form E, 10 pp., Physical Science, Form E, 10 pp.

These tests are the 1939 examinations used by the Chicago City Junior Colleges. The English qualifying examination is a three-hour test containing both essay and objective exercises, the latter being designed for machine scoring. The essay sections of the tests used in the survey courses require one hour and contain a variety of exercises calling for responses in the form of sentences, paragraphs, or brief essays. Each objective section requires two hours of testing time, consists of several types of exercises, and may be scored on an electric scoring machine. The English placement test has no time limit, is objective, and machine scored. The exemption tests in biological science and physical science require one hour and twenty minutes and are objective.

GATES, ARTHUR I. *Gates Reading Readiness Tests* (New York: Bureau of Publications, Teachers College, Columbia University, 1939), 7 pp.

The Gates tests were constructed to measure readiness for beginning reading, to predict the rate of development of reading ability, and to diagnose the pupil's status and thus reveal his needs in each of several of the most important abilities required in learning to read.

MALLER, J. B., and GLASER, EDWARD M. *Interest-Values Inventory* (New York: Bureau of Publications, Teachers College, Columbia University, 1939), 8 pp.

This test is designed to measure the relative dominance of four major types of interest or basic values within the individual: theoretic, aesthetic, social, and economic. It is for use with high school and college students and adults.

SHAFFER, LAURANCE F. *Objective Tests for the Psychology of Adjustment* (Chicago: Houghton Mifflin Company, 1939), 17 pp.

These tests consist of one page each with twenty to twenty-five objective items for each of the seventeen chapters of Dr. Schaffer's text and are designed for use with college classes using the text.

STRANG, RUTH. *Examiner's Reading Diagnostic Record for High School and College Students* (New York: Bureau of Publications, Teachers College, Columbia University, 1939), 20 pp.

This reading diagnostic record form includes identifying data, summary of results of standardized tests, summary of scholastic achievement, medical examination and reports on physical condition, developmental and educational history, present reading interests, other interests and activities, present reading status (tests of oral reading and use of dictionary included in the record form), summary, recommendations, and follow up.

RESEARCH NEWS and COMMUNICATIONS

Address all research news and communications to Carter V. Good, Teachers College, University of Cincinnati, Cincinnati, Ohio.

Yearbook of the John Dewey Society.—*Teachers for Democracy* is the tentative title of the Fourth Yearbook of the John Dewey Society, a commission of the Progressive Education Association. The yearbook, to be published by D. Appleton-Century, will probably be officially released in January. The book will deal with the problems of teacher education and is written from the viewpoint of suggesting solutions to the problems of preparing teachers who will have the qualifications necessary for democratic education.

The first chapter, by Frank E. Baker, president of the Milwaukee State Teachers College, and E. O. Melby, dean of the School of Education, Northwestern University, describes the present status of teacher education and the resultant problems. The second chapter, by William W. Wattenberg, of the Chicago Teachers College, sets forth the tasks before the profession as these find expression in work confronting institutions preparing teachers. This is followed by a description of the qualifications needed by teachers in democracy's school, by G. Robert Koopman, of the Michigan State Department of Public Instruction. E. T. McSwain, of Northwestern University, contributes a chapter on conceptions of the learning process and the social implications for teacher education of the various conceptions.

The life and program of teacher-education institutions, as these affect the atmosphere of the institutions and influence the attitudes and social abilities of students, are described by Nila Banton Smith, of Indiana University, and Frances Martin, of Central State Teachers College, Mount Pleasant, Michigan. Frank E. Baker analyzes the administrative factors necessary for intra-institution democracy. The problems of general education are set forth in a chapter by Donald Cotrell, of Teachers College, Columbia University. Whit Brogan, of Northwestern University, treats of the considerations influencing the selection of professional subject matter. This chapter is followed by one describing the treatment of the materials of professional education and the integration of theory and practice in an appropriate student-teaching program, by John J. DeBoer, director of practice teaching at Chicago Teachers College. The basic problems and procedures of in-service continuing education are set forth by Walter Anderson, of Northwestern University. This is followed by a chapter on the education of college teachers by Donald Cotrell. The last chapters of the book, which consider the basic social context and social implementation of the teaching profession, are written by George E. Axtelle, of Northwestern University.

The yearbook is edited jointly by George E. Axtelle and William W. Wattenberg.

Social Studies Yearbook.—The Tenth Yearbook of the National Council for

the Social Studies is on the subject: "The In-Service Growth of Social Studies Teachers." It will be distributed to members of the National Council soon after November 1 and will be presented and discussed at the annual meeting of the N. C. S. S. in Kansas City, November 24 and 25.

"In-service growth," rather than "in-service training," was chosen as the subject, because the editor and the committee on publications wished to make a direct appeal to the teacher. The Yearbook should also be of interest to those responsible for the in-service training and guidance of teachers.

The problems of the beginning teacher of the social studies are discussed by James A. Michener of the Harvard Graduate School of Education. Lewis W. Williams, University of Illinois, contributes a chapter on problems of placement, promotion, and tenure. A reading plan for social studies teachers is outlined by Edgar B. Wesley of the University of Minnesota. A. L. Kerbow of the University of Houston discusses the problem of graduate study in its relation to in-service growth, basing his conclusions on a study just completed in Houston. Workshops, curriculum laboratories, and courses differing from the traditional methods courses are described by Anna B. Peck of the University School at Lexington, Kentucky. Edward G. Olsen of Colgate University presents a study of tours and travel courses for social studies teachers. Laura F. Ullrick of the new Trier Township High School, Winnetka, Illinois, and Charles H. Coleman of Eastern Illinois State Teachers College, have collaborated in preparing a chapter on the community contacts of the teacher in relation to his professional growth. Experimental teaching as a means of professional growth is discussed by J. G. Umstattd of the University of Texas. The possibilities of in-service growth through membership and active participation in professional organizations are described by Edith B. Ongley, Director of Social Studies, Binghamton, New York.

The Committee on Publications of the National Council for the Social Studies has as its members Wilbur F. Murra and James A. Michener, both of the Harvard Graduate School of Education, and Fremont P. Wirth of the George Peabody College for Teachers. The Editor of the Tenth Yearbook is Burr W. Phillips, Head of the Social Studies Department, Wisconsin High School, and Associate Professor in the Teaching of History, University of Wisconsin, who forwarded the present statement.

Teacher Education.—The Commission on Teacher Education of the American Council on Education, formed in 1938 for a five-year period, with Karl W. Bigelow as director, launched during 1939 a "Coöperative Study of Teacher Education" involving 20 higher institutions and 14 school systems and groups of school systems. The major purposes of this study are to encourage a rapid translation into practice of generally accepted knowledge in teacher education and to stimulate experimentation, broadly conceived, with programs of teacher education and teacher growth in service. It has been assumed that these two purposes would be best served by working closely at the outset with a limited number of institutions and school systems.

The units[1] selected will be assisted in

[1] *Universities*—Columbia University (consisting of Columbia, Barnard, and Teachers Colleges), Ohio State University, Stanford University, University of Nebraska, University of North Carolina, University of Texas; *Liberal Arts Colleges*—Claremont Colleges, College of St. Catherine, College of William and Mary, Middlebury College, Oberlin College; *State Teachers Colleges*—Colorado State College of Education, Greeley; Eastern Kentucky State

a thorough examination of their problems and their curricula and will be helped in various ways to become, or to continue to serve, as demonstration centers of best practice in the education and professional development of teachers both prior to and subsequent to employment. In addition, the Commission will endeavor to disseminate as rapidly as possible to the educational world information concerning promising developments in teacher education wherever these may be found.

In August, 1939, a unique conference attended by approximately 125 persons, most of them representatives of the units engaged in the Coöperative Study of Teacher Education, was held at Bennington College, Vermont. The central objective of the conference was the planning of coöperative activities for the coming 3½ years. The delegates constituted a cross-section of American education, including persons of diverse backgrounds and special interests who were not accustomed to conferring together. There were present representatives from universities, liberal arts colleges, and teachers colleges as well as from rural, urban, and suburban school systems. These included educationists and subject-matter specialists, college presidents, deans, and professors, superintendents of schools, principals, supervisors, and classroom teachers. At the close of the two-week conference period, these individuals, as a result of democratic discussion, had reached working agreements as to the major problems of teacher education and how their institutions working jointly with the Commission might best attack these problems in order to make the greatest possible contribution to the improvement of American education. The conference was conducted in such a way as to ensure the participation of each member at every stage. It was agreed that reports of this and possible later conferences, and of the work in the individual institutions, should be made generally available to the profession.

Plans for certain studies involving entire states and for activities in collaboration with educational organizations are now being formulated by the Commission.

School Plant Research.—Following a series of conferences sponsored by the American Council on Education, which began in November, 1935, the School Plant Research Council was organized as one of the standing committees of the American Council on Education. During these early meetings extended consideration was given to the need for research in the school building field and if such a need existed, how a research program should be organized, financed, and administered. Growing out of these conferences there has been developed a bulletin entitled, "School Buildings and Equipment," which sets forth the specific areas where the need for investigation is most evident and the proposed organization to carry on those investigations and to dis-

Teachers College, Richmond; New Jersey State Teachers College, Newark; Southern Illinois State Normal University, Carbondale; State Teachers College, Milwaukee, Wisconsin; State Teachers College, Troy, Alabama; Western State Teachers College, Kalamazoo, Michigan; *Negro Colleges*—Prairie View State College, Texas; Tuskegee Normal and Industrial Institute, Alabama; *School Systems*—Bedford County, Virginia; Central District, Caledonia County, Vermont; Colquitt County, Georgia (consisting of the county and Moultrie city school systems); Denver, Colorado; Des Moines, Iowa; Detroit, Michigan; Greenville county, South Carolina (consisting of the county, city, and Parker District systems, with the Greenville County Council for Community Development and Furman University coöperating); Houston, Texas; Los Angeles County, California (consisting of the county, city, Pasadena, Santa Monica, and Burbank systems); Newton, Massachusetts; New Trier Township, Illinois (consisting of the township high school system and the elementary systems of Glencoe, Kenilworth, Wilmette, and Winnetka); Norris, Tennessee; Philadelphia, Pennsylvania; and Spokane, Washington.

seminate the findings. This statement was published in April, 1939, by the American Council on Education.

The first meeting of this Research Council was held in Washington on January 28 and 29, 1938. This was followed by meetings in May and December, 1939, and in March, 1939. Inasmuch as the Research Council was newly organized on the one hand, and financial support for carrying on investigations has not yet been received, these four meetings have been largly devoted to the following considerations:

1. Preparation of the manuscript referred to above and consideration of other similar manuscript which might be prepared in specific areas such as lighting, heating, and ventilating.
2. Possibilities of financial support for the program as outlined by the Council.
3. Preparation of detailed outlines for investigations in the fields of lighting, toilet facilities, and school equipment.
4. Other problems in the field which might be similarly prepared for study.
5. Discussion with the representatives of agencies interested in research in the school building field. These agencies are the United States Office of Education, United States Public Health Service, Bureau of Standards, Public Works Administration, and the National Research Council.

Two of the three studies listed under Item 3 above, toilet fixtures and school equipment, are now actually under way. The status of each of these is as follows:

Toilet Fixtures.—Workers in the school plant field have long believed that the commonly accepted standards on the number of fixtures required are too high. This belief has been supported by a few local studies which have been made. Therefore the Council has under way a nation-wide study of this problem to see what the facts are. The reason for making it nation-wide is to ascertain the influence, if any, of climate on the number of toilet fixtures needed for a given school enrollment. Furthermore, the study is being carried on in both cities and rural districts in schools of different sizes and internal organizations and in schools of different racial makeup. The plan is to make an actual "Time-Use Study of All Toilet Facilities" in a school for one entire day in February and again in May.[1] The first of these has already been made. This study is under the immediate direction of Francis R. Scherer, Superintendent of School Buildings in Rochester, New York, and a member of the Council. Printing costs and other incidental expenses are being borne by the American Council on Education.

School Equipment.—From ten to fifteen percent of the cost of school buildings is devoted to equipment. On that basis, therefore, public funds to the extent of approximately $45,000,000 were spent on this item alone in 1930. For the alternate years of 1920-1936 inclusive the expenditure for school equipment alone amounted to about $300,000,000. Despite these large expenditures, little headway has been made in the development of quality specifications which would classify a given item of equipment according to certain definite requirements to be incorporated in these specifications; for example, specifications which would group fixed seats into probably three classes according to quality. Thus a board of education which wished to limit the

[1] Because some of the data from the February count were not usable, it was found necessary to postpone the May count until the fall of 1939-40.

bids for fixed seats to those regarded most satisfactory would so indicate in their notice to bidders. Competition would then be based on comparable quality of product.

At the December, 1938, and March, 1939, meetings of the Council, representatives of the equipment industry were invited to be present to discuss whether the plan in mind was a feasible one. A similar discussion was also held at the Cleveland meeting of the A.A.S.A. These representatives were in general accord that such a plan, if thoroughly done, would be highly desirable. This study, therefore, is going forward under the immediate direction of Raymond V. Long, Director of the Division of School Housing in Virginia and a member of the Council.

In conclusion, members of the Council are of the belief that definite headway has been made since its organization. Furthermore, they are confident that the need for research is so evident in school buildings, which annually require approximately one-third of the total cost of public elementary and secondary education for their construction, operation, and maintenance, that financial support for the Council's program will be forthcoming.

This information was provided by T. C. Holy of Ohio State University, a member of the committee in charge of the project.

Secondary School Standards. — After six years of intensive work (research, experimentation, analysis, revision, and refinement) the Coöperative Study of Secondary School Standards has completed its work. It now has ready to offer to the secondary schools of the country a body of materials and procedures for evaluation of secondary schools which are believed to be more valid, more flexible, and more stimulating to improvement than any available in the past.

A sum of approximately $200,000 has been expended by the Coöperative Study in the past four years in the development and refinement of these materials. They were first carefully tested in 200 secondary schools of all sizes and types in all parts of the country. As a result of this tryout the materials were extensively revised and published as "1938 editions" and again tested in 90 additional schools. During the spring and summer of 1939 they have been further revised and rewritten. The "1940 editions," described below, represent the best judgment and experience of the Coöperative Study. It is expected that further revisions will not be made for at least five years.

1. Evaluation of Secondary Schools: General Report.—A comprehensive report of the purposes, history, methods, and recommendations of the six-year study. Essential for any one who wishes to know the background and reasons for the Coöperative Study's results and recommended procedures.

2. Evaluation of Secondary Schools: Supplementary Reprints.—A bound volume of 29 reprints of articles previously published in various educational journals, each reporting on some phase of the Coöperative Study. Intended as a companion volume for the *General Report* listed above.

3. How to Evaluate a Secondary School.—A manual of instructions designed to accompany *Evaluative Criteria* and *Educational Temperatures.* Gives all information and instructions necessary for actual use of the procedures recommended by the Coöperative Study for evaluation of a school. Contains chapters of instructions and suggestions to schools, to visiting committees, and to

supervisory and accrediting agencies. Includes, as an appendix, the charts from *Educational Temperatures*, filled out for five representative schools.

4. Evaluative Criteria.—Contains approximately 1600 checklist items and 500 evaluations covering all significant phases of the modern secondary school. Includes the following sections (also published as separate pamphlets): Philosophy and Objectives, Pupil Population and School Community, Curriculum and Courses of Study, Pupil Activity Program, Library Service, Guidance Service, Instruction, Outcomes of the Educational Program, School Staff, School Plant, School Administration, Data for Individual Staff Members, and Summary Forms.

5. Educational Temperatures.—A series of charts for exhibiting in vivid and graphic form the results of an evaluation made by means of the material in the *Evaluative Criteria*. Shows relative standing or "educational temperatures" of 110 significant aspects, so presented as to facilitate comparisons with schools of different size, type, accreditation status, and regional location, or with schools of all sizes and types in all parts of the country.

Commission on the Secondary School Curriculum.—The period in which this Commission of the Progressive Education Association was to work ended on July 1. Consequently the Commission is engaged in sending to the publishers reports of the various committtees. On September 15, *Reorganizing Secondary Education*, by V. T. Thayer, Caroline B. Zachry, and Ruth Kotinsky appeared. This attempts to state the fundamental point of view that has permeated the work of the Commission and its committees.

The reports which have been published and those scheduled for publication are listed separately. The following books have already been published by the D. Appleton-Century Company:

Lawrence H. Conrad, with the Creative Committee, *Teaching Creative Writing* (1937)

Committee on the Function of Science in General Education, *Science in General Education* (1938)

The following books are scheduled for publication by the D. Appleton-Century Company in 1940 (all titles are tentative):

Committee on the Function of Art in General Education, *Art in General Education.*

Committee on the Function of English in General Education, *Language in General Education.*

Committee on the Function of Mathematics in General Education, *Mathematics in General Education.*

Committee on the Function of the Social Studies in General Education, *The Social Studies in General Education.*

Elbert Lenrow, for the Committee on the Function of English in General Education, *Guide to the Reading of Fiction in General Education: Bibliographies of 1,500 Novels Selected, Classified and Annotated for Use in Meeting the Needs of Individuals in Senior High School and College.*

Discussion Group Project.—The National Association of Secondary School Principals, acting through its Planning Committee, has developed a long-range program for the improvement of the American high schools. The program is being promoted through the democratic process of discussion and coöperative thinking. The following steps have been taken or are being taken:

1. A study of the issues and objectives of secondary education by a body of em-

nent educators—The Orientation Committee—followed by a statement of conclusions, published as "The Issues of Secondary Education" and the "Functions of Secondary Education."

2. The sustained and systematic discussion of these issues and objectives by the high school principals and teachers of the nation. In order that this exercise in coöperative thinking might be effectively carried out, discussion group machinery was set up. Each state was organized. Discussion guides were distributed and discussion techniques were studied. Thousands of principals and teachers participated in this organized discussion. Probably never before in the history of education have so many educators fixed their attention upon any educational program.

3. The next step is from theory to practice. Whenever an announced objective of high school education is set forth, the questions arise, "Where is this objective being realized in actual practice?" "What school or schools now follow this plan?" "What can we learn from actual experience?" The proposal now is to find answers to these questions and to lay the results before discussion groups through the nation for their consideration.

4. An appraisal of the most promising practices in the various fields of secondary education, of the practices which point the way to the solution of the outstanding problems of the high schools, with suggested programs of action, is the work of the Implementation Committee which has been appointed by the National Association of Secondary-School Principals.

Walter E. Myer, Washington, D. C., is the national director of the project.

Research Division, N. E. A.—Current projects are:

1. Research Bulletins
   a. "Organizations Interested in International Relations," September. Annotated list of agencies working for peace, education, and other goals.
   b. Teacher Load, November. Nearly 4000 teachers tell why and where their assignments are most heavy. Suggestions are made for improving conditions.
   c. Salaries as Related to Training and Experiences, January. Summary of the general patterns of training and experience found in city school systems; comparisons made with salaries paid in 1938-39.
   d. Other topics under study and publication in March or May, 1940, or at later date: Status of the Teaching Profession (a summary of the characteristics and status of teachers in the United States); The Gifted Child (what is being done for those with rare talents); The Problems of Rural Education (some trends and questions); The Distribution of Supervisory Authority (trends in assignment of supervisory power in city school systems).
2. Yearbooks
   a. 1940 yearbook, American Association of School Administrators, "Safety Education." Present practices and future possibilities of safety instruction in schools. (Available February, 1940.)
   b. 1941 yearbook, American Association of School Administrators, "Education for Family Life." Preparation during 1940 and publication in February, 1941.
   c. 1940 yearbook, Department of Elementary School Principals, "Meeting Special Needs of the Individual

School Child." Available about September 1, 1940.

3. Committees
   a. Tenure. Several studies are planned dealing with tenure in private liberal arts colleges, professional training of teachers with and without permanent tenure status, proportion of married women teachers, and court decisions in 1939. Available by June, 1940.
   b. Supply, Preparation, and Certification of Teachers. Preparation of handbook on how to make a statewide study of supply and demand and revision of tentative principles of teacher education. Available in spring, 1940.
   c. International Relations. A sample study of what is being done in high schools to develop an interest in international relations. Available in June, 1940.
   d. Credit Unions. Handbook now being prepared on how to organize teacher's credit union. Available late in 1939.
   e. Tax Education. Preparation during the year of materials on taxation for high-school classes. Available in spring, 1940.
   f. Equal Opportunity. Prepare a bibliography of books on negroes acceptable to members of that race. Available in 1940.
   g. Coöperatives. Edit the 1939 report dealing with the history and purpose of credit unions and units of work for use in schools. Available late in 1939.
   h. Certification of Superintendents. Help complete case studies of successful superintendents for American Association of School Administrators. Probably available by winter meeting, 1940.
   i. Teacher Education. Several studies may develop in connection with program of American Association of Teachers Colleges.
   j. Retirement. Several reports in progress dealing with court cases concerning retirement, interests of retired teachers, and status of local retirement systems. Available during 1939-40.
4. Miscellaneous
   a. Safety education projects include an analysis of safety education courses, a school building checklist for principals, summary of research in safety, and units of work in teaching safety. Available during the year, 1939-40.
   b. Salary Scheduling. A series of bulletins on how to draft a salary schedule. Now in mimeographed form; being revised for printing.
   c. P.-T.A. book. A review of public education for laymen being edited for the National Congress of Parents and Teachers. Available about November, 1939.
   d. Tax Distribution. Possible study of the distribution of state tax funds with special reference to welfare, relief, and pensions.

Citizenship Education.—To implement the entire program of the Educational Policies Commission, even within the single field of civic responsibility, is a task for which American education at the moment lacks adequate experience. A beginning is being made with one highly critical and urgent part of the total area of civic responsibility. With the aid of a special grant from the General Education Board this phase of the work began on July 1, 1939, and will continue for one year.

The purpose of the project is to improve the effectiveness with which American schools develop in young people an

intelligent, appreciative, and active loyalty to democracy. This outcome involves the application of the principles, processes, and values of democracy to the affairs of everyday life.

Civic responsibility as a major purpose of American education is now rather widely accepted in theory, both by the profession and the lay public. Practice, however, lags discouragingly behind the accepted philosophy. Appropriate working methods and materials are not yet developed and made widely available. The constructive programs of a few forward-looking school systems are not known outside their immediate vicinities. Sporadic efforts of individual teachers and supervisors are engulfed by the inertia and diffused effort of the huge educational enterprise.

"Civic education" is used as a convenient brief term to refer to the project as described above. Efforts will be confined largely to the elementary and secondary schools, with emphasis upon the latter. One requirement of the project will be a clear-cut description of the desired educational outcomes. What are the ingredients of loyalty to democracy?

It has proved to be convenient in planning this project to identify certain steps which will probably occur in some such sequence as the following:

1. To describe (in the light of previous publications of the Educational Policies Commission and all other useful and available sources) the essential ideals and conditions of democratic life.

2. To identify, in terms of the above specifications, at least some of the features which appear to characterize a good educational program for developing an appreciative, intelligent, and active loyalty to democracy.

3. To bring together in outline form all of the practicable methods for estimating the success of such a civic education program.

4. To make a fairly intensive first-hand analysis of forty or fifty selected schools or school systems which appear to have effective programs of civic education. These school systems would be studied against the criteria developed under item 2 above, and by using the methods developed under item 3.

5. To describe, in the form of a "case-book" of successful practice in civic education, the philosophy, nature, and methods of the best of these programs, together with the evidence for believing that they are actually effective.

6. To exert leadership, within the school systems studied and elsewhere, with a view to developing new and improved civic education techniques.

7. To draft a tentative program for conducting nationwide demonstrations or other projects in the field of civic education which will extend more widely the superior educational practices which have been discovered and described.

The staff members of the Civic Education Project include:

G. L. Maxwell, Assistant Secretary, on leave from the University of Denver, and for the last three years, Assistant Director, Education Program, Works Projects Administration, Washington;

Howard E. Wilson, Specialist (Social Studies) on leave from Harvard University;

Samuel Everett, Specialist (Community Relations), on leave from Northwestern University;

Oliver H. Bimson, Specialist (Administration), on leave from the Assistant Superintendency of Schools, Lincoln, Nebraska;

Harold Benjamin, Consultant, part time, Dean, School of Education, University of Maryland.

Office of Education Transferred.—In accordance with the President's First Plan on Government Reorganization, the Office of Education has been transferred from the Department of the Interior to the Federal Security Agency. The transfer, according to the joint resolution which passed both Houses, became effective July 1, 1939. The organizations, in addition to the Office of Education, which compose the new Federal Security Agency are: Social Security Board, formerly an independent establishment; the United States Employment Service, from the Department of Labor; the Public Health Service, from the Treasury Department; the National Youth Administration, from the Works Progress Administration; and the Civilian Conservation Corps, formerly an independent agency.

The President's Second Plan also affects the Office of Education in that it transfers the Radio and Motion Picture Divisions of the National Emergency Council to the Office.

In the message from the President, transmitting the Reorganization Plan to Congress, he stated that: "Because of the relationship of the educational opportunities of the country to the security of its individual citizens, the Office of Education with all of its functions, including, of course, its administration of Federal-State programs of vocational education, is transferred from the Department of the Interior to the Federal Security Agency. This transfer does not increase or extend the activities of the Federal Government in respect to education, but does move the existing activities into a grouping where the work may be carried on more efficiently and expeditiously, and where coördination and the elimination of overlapping may be accomplished. The Office of Education has no relationship to the other functions of the Department of the Interior."

Comparative Education.—The International Institute of Teachers College, Columbia University, is now continuing with only one project—the *Educational Yearbook*. The volume for 1939, now in the press, is devoted to a series of articles on "The Meaning of a Liberal Education." Contributions will be included from seventeen countries, as follows: Argentine Republic, Australia, Brazil, Chile, Denmark, England, France, Germany (1) pre-Nazi, and (2) Nazi, Holland, India, Italy, New Zealand, Norway, Scotland, Sweden, Union of South Africa, and the United States. In order to obtain as diversified views as possible three articles have been prepared for the United States, three for England, and two for France. Since the issue is one of crucial importance in this country, it is of interest to note the continued faith in the traditional concept of a liberal education in some countries, and the plans for readaptation in others. The differences are based on differences in philosophy of education and psychology in some cases, in others on the socio-economic demands of contemporary life.

The topic for the *Educational Yearbook, 1940*, is "Adult Education," on which contributors from seventeen countries have been invited to write. A number of the articles have already been received. The volume should appear in the fall of 1940.

The international study of examinations is also still under way, and a few reports yet remain to be completed in England and the Scandinavian countries before a final summary volume can be prepared.

This progress report was forwarded by I. L. Kandel, editor of the Yearbook.

Columbia Studies.—Projects of the Division of Psychology, Institute of Educational Research, Teachers College, Columbia University, are as follows:

E. L. Thorndike is completing a study of various matters concerning the psychology of economics, business, government, law, and social work. A large volume will appear shortly under the title of "Human Nature and the Social Order," and certain special researches will be published from time to time in appropriate journals. This will include studies of business inventions, philanthropic inventions, wage ratios, 100 Italian cities following a general plan of the report in "Your City," and the press in American cities.

Irving D. Lorge and Dr. Thorndike have completed the "Semantic Count of English Words," comprising a count of approximately four-and-a-half million words classified by the frequency of the occurrence by their separate meanings. Dr. Lorge is continuing a survey of "Metropolitan Youth" in order to ascertain the relationship between various measures of intellectual, clerical, and mechanical ability and subsequent vocational success and adjustment.

Division of Educational Research, Philadelphia.—In an effort further to improve the guidance activities in the senior high and vocational schools of Philadelphia, the principals and guidance personnel decided upon a thorough-going survey of guidance procedures to be carried on during the current term. As a basis for this survey the principals and counselors agreed to conduct a complete self-evaluation of each school's practices by the use of Blank G, Guidance Service, of the Evaluative Criteria of the Coöperative Study of Secondary School Standards. These procedures of evaluation were applied recently in one Philadelphia senior high school in coöperation with the Pennsylvania State Department of Public Instruction. In this case the Evaluating Committee consisted of high school principals and department heads of other schools, both public and parochial, members of the staff of the Division of Educational Research, high school principals and superintendents of nearby counties, and members of the departments of secondary education of nearby universities.

The present self-evaluation will be conducted in each school by the guidance personnel of that school. Further study of these evaluations will be the responsibility of a joint committee representing all the schools. In this way it is planned to secure a clear picture of the variety and extent of guidance problems from which the guidance counselors may set up procedures and organizations for adequate guidance service.

The type of evaluation sponsored by the Coöperative Study of Secondary School Standards is applicable to all aspects of secondary school work. Perhaps the experience with the use of these criteria in the field of guidance may lead to extension of coöperative studies to other areas of secondary school activity.

For the year ended June, 1939, the testing program was extended by offering a wide selection of tests for grades 7 to 12 in twenty-two subjects. This type of testing program has important advantages in that the schools may select tests especially suitable for their needs. For example, in a school in which most of the pupils are preparing for college entrance, the tests selected may emphasize academic subject matter in the upper grades. In a school in which most of the pupils are in low IQ intervals the tests may be directed more toward study of skills in reading, English usage, and fundamental operations of arithmetic. These records

cumulated over the school life of a pupil offer suggestions to guidance counselors which are sometimes of great importance.

One of the most difficult obstacles to overcome in the extension of test resources is the amount of work involved in scoring tests and making records of the results. In June, 1939, mental ability tests in grade 9A were partially scored by re-running pupil answer sheets through the printing press. By this method it was possible to score tests in approximately one-half the time formerly required.

Changes in the distributions of IQ's of high school pupils are noted for 1938-39. These are the result largely of: (1) the increased provision for pupils of low IQ in the modified and prevocational curriculums, (2) the increased enrollments due to the extension of compulsory education to include sixteen-year-old pupils, and (3) the organization of new high schools. Pupils for the prevocational and modified curriculums in senior high school have apparently been selected from those who formerly elected the academic curriculum. In junior high school, pupils in the prevocational curriculum are those who formerly elected any other curriculum. Approximately 1000 pupils have been added to grade 10A, and 1500 pupils to grade 9B, as a result of the extension of compulsory education to include sixteen-year-old pupils. The resulting increase has brought about a reduction in the median IQ of pupils in grades 9 and 10 from 106 to 105.

Test achievement in the Philadelphia Junior Test in English Usage, as in other Philadelphia tests, is interpreted for each pupil in terms of his IQ. This is provided by a table printed on the test and by a summary of relative scores and expectancies (X scores) prepared for each class. Pupils who do not attain relative scores within one point of their X scores are exceptional. Such cases throughout the city amounted to only 4 percent of total of 10,407 pupils. Class and sch summaries provide a direct comparison the mean expectancy (X score) and mean relative score. The difference tween these two measures in most schools may be considered negligible. Pupils with high expectancies, X scores of 4 and in schools with high average X scores achieve higher in relation to their expectancy than do pupils with similar high expectancies in other schools. A similar finding was also found for scores in reading tests.

Promotion rates in junior and senior high schools remain substantially the same as those reported in previous years. Continued study of promotion rates has shown wide application of negative guidance through failure of pupils. Repetition of subjects made necessary by the requirements for graduation has resulted too frequently in ineffective achievement. As a result of new standards for graduation the high school program of studies has been made more flexible. This flexibility should result in better pupil adjustment and a more satisfactory rate of progress in the senior high schools.

The Director of the Division of Educational Research is Philip A. Boyer.

Studies in Early Graduate Education.—What were the chief factors in the successes of Johns Hopkins University, Clark University, and the University of Chicago in their early days, and to what extent can American universities of today learn from them? Does the experimentation of these three institutions, as well as of Harvard, Yale, Princeton, Columbia, the University of Michigan, and other pioneers in graduate work, throw any light on present problems of university education, especially at the graduate level? These are the questions considered in Bulletin Thirty of the Carnegie Foundation for the Advance

ment of Teaching, *Studies in Early Graduate Education*, prepared by W. Carson Ryan, staff associate of the Foundation.

The unusual achievement of the early days at Hopkins, Clark and Chicago—established in 1876, 1888, and 1893—came primarily because those in charge of these universities recognized the urgency of certain contemporary needs that were not being met by the conventional college education of the day; because, moreover, in striving to meet these needs those in authority had the wisdom to place the human element in the educational process—the students and faculty—ahead of buildings, enrollments, administration, and organization; and because the men who headed the universities were themselves vigorous personalities who insisted upon the highest possible quality of intellect and character in the institutions they directed.

This concern for careful selection of faculty and students was accompanied in all three institutions by an insistence upon conditions of teaching and learning that encouraged independent investigation, intimate and informal teacher-student relationship, and a minimum of administrative machinery. At the end of the first fifteen years President Gilman was convinced that what had brought students to Johns Hopkins, among other things, was the belief that "a strong personal interest would be shown in every student," and at Clark University, Hall and his associates found the highest possible intellectual aims wholly compatible with friendliness and sympathetic concern for the individual.

Freedom of teaching and learning was inherent in these pioneer beginnings at Hopkins, Clark and Chicago. It would be difficult, for example, to find any stronger position on behalf of university freedom than that taken by President William Rainey Harper of the University of Chicago. A faculty composed of individuals of widely differing views had been brought together at Chicago and were expected to be independent in their research and free in their teaching. The principle of "complete freedom of speech on all subjects" was insisted upon from the first, and President Harper stated as explicitly as he could that no donor had the right to interfere with the teaching in the university. Even though freedom of speech might at times be abused, he said, this was no justification for curtailing it—"the abuse of it is not so great an evil as the restriction of such liberty."

Educational Research Through National Agencies.—A majority of the English speaking countries now possess some recognized national center which exists for the purpose of conducting educational inquiries and of supplying information on educational developments. New Zealand has a council which, like the Australian Council for Educational Research, is an independent agency subsidized by the Carnegie Corporation of New York. The United States has a federal Office of Education. South Africa possesses a National Bureau of Educational and Social Research maintained by the Union Government. Scotland has a Council for Educational Research formed and maintained through the coöperation of the Educational Institute of Scotland and local education authorities. It is interesting to learn that a meeting of educationists was held in Toronto in January of this year to discuss the possibility of establishing a Canadian Council for Educational Research.

Review of Australian Education in 1938.—The Australian Council for Educational Research has printed the first number of what it is hoped will be an

annual survey of developments in Australian education. Practically all branches of education are dealt with in some 200 pages. It has been considered appropriate that a relatively large section of the space in the first number should be taken up with a general account of the structure of the six State educational systems. A specially contributed article deals with the Catholic school system. Using this general statement as a background the greater portion of the book is devoted to an account of various educational developments which have taken place during the last year or so. A critical review of these developments is presented in a concluding chapter. For purposes of reference there is included an extensive bibliography of significant books and articles on Australian education.

The Council has published *The Standardization of Intelligence Tests in Australia*, by G. A. McIntyre. This is the account of the standardization of two intelligence tests applied to 30,000 children throughout Australia in March, 1936. Two tests were used, a non-verbal test and the Otis Intermediate test revised to suit Australian conditions, and full information is given in regard to the preparation of the tests and the results obtained. Since the Otis test is widely used in America, and the New Zealand Council for Educational Research gave the tests at the same time they were applied in Australia, it has been possible to make international comparisons, in addition to the more usual interstate, urban, rural, and sex comparisons. A survey of previous work in Australia gives a valuable history of intelligence testing in Australia, with a comparison of present and past findings.

New Zealand Council for Educational Research.—The Council has been informed that the original grant made by the Carnegie Corporation for a period of five years will be continued for another five years. It is fully anticipated that during the second five-year period additional sources of finance will be provided within New Zealand and that when the experimental years are over the results will insure that research will find a permanent place in the system of education. The Council has published a total of 12 reports and studies.

School Surveys. — Comprehensive surveys have recently been made of the St. Louis, Missouri, and Stockton, California, schools. The former was directed by George D. Strayer of Columbia University, with the assistance of numerous colleagues. The report covers 494 pages. The 616 page report of the Stockton system was prepared by Jesse B. Sears and associates of Stanford University, with some other assistance.

JOURN

Volume XXXIII

A QUALITA

Editor's no
ceived more att
Many important
and research.
information rela
experience that

ACHIEVEMENT
ing is the major
upon social valu
than the produc
aspects of reading
with respect to q
means of objecti

As long as the
of objective evalu
ment, the appra
must remain larg
be made possible
man judgment.
provided, individ
students may be
mature reading,
The fabrication
quantitative fact

Symonds[2] has
standardization a
dependence upo

[1] National Coun
(New York: D. Ap
[2] Symonds, Per
Century Company,